

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 17:34:18 ; Search time 17970.8 Seconds
(without alignments)
-1569.817 Million cell updates/sec

Title: US-08-852-495c-2_COPY_1_29000
Perfect score: 29000
Sequence: 1 CACACACACACACACACA.....AACCTGTGCTCTCTGGGTTTC 29000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 882769 seqs, -486395729 residues
Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_cm:*
- 4: gb_ov:*
- 5: gb_pat:*
- 6: gb_ph:*
- 7: gb_pl1:*
- 8: gb_pl2:*
- 9: gb_pr1:*
- 10: gb_pr2:*
- 11: gb_pr3:*
- 12: gb_ro:*
- 13: gb_sts:*
- 14: gb_sy:*
- 15: gb_un:*
- 16: gb_vl:*
- 17: em_fun:*
- 18: em_hum1:*
- 19: em_hum2:*
- 20: em_in:*
- 21: em_om:*
- 22: em_or:*
- 23: em_ov:*
- 24: em_pat:*
- 25: em_ph:*
- 26: em_pl:*
- 27: em_ro:*
- 28: em_sts:*
- 29: em_sy:*
- 30: em_un:*
- 31: em_vl:*
- 32: gb_htg1:*
- 33: gb_htg2:*
- 34: gb_in1:*
- 35: gb_in2:*
- 36: em_ba1:*
- 37: em_ba2:*
- 38: em_hum3:*
- 39: em_hum4:*
- 40: gb_pr4:*
- 41: gb_htg3:*
- 42: gb_htg4:*
- 43: gb_htg5:*
- 44: gb_htg6:*

- 45: gb_htg7:*
- 46: em_htg1:*
- 47: em_htg2:*
- 48: em_htg3:*
- 49: em_hum5:*
- 50: gb_pl3:*
- 51: gb_pr5:*
- 52: gb_htg8:*
- 53: gb_htg9:*
- 54: gb_htg10:*
- 55: gb_htg11:*
- 56: gb_htg12:*
- 57: gb_htg13:*
- 58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
C 1	87.2	0.3	103	9	HUMALCE221
C 2	87	0.3	107	9	HUMALCE162
C 3	87.2	0.3	108	10	HSLDLRN2
C 4	83.6	0.3	108	10	HSLDLRN2
C 5	79.8	0.3	108	10	HSLDLRD1
C 6	79.8	0.3	108	10	HSLDLRD2
C 7	80	0.3	108	11	HSU67803
C 8	75	0.3	103	9	HUMALCE221
C 9	75.2	0.3	108	11	HSU67804
C 10	74.2	0.3	108	9	HUMDID03M5
C 11	73.6	0.3	108	10	HSLDLI12
C 12	73.4	0.3	110	11	HSU67807
C 13	73	0.3	103	13	HS8IC8R
C 14	73	0.3	108	10	HSLDLRD1
C 15	73	0.3	108	10	HSLDLRD2
C 16	72.4	0.2	101	10	S79560
C 17	71.6	0.2	94	9	HUMHGAL
C 18	70.8	0.2	90	9	HUMDLRLFL
C 19	70.8	0.2	91	13	HUMUT8164A
C 20	69.8	0.2	108	13	G32614
C 21	69.8	0.2	110	9	HUMALCE43
C 22	68.8	0.2	106	13	G32743
C 23	69	0.2	108	11	HSU67803
C 24	68.8	0.2	108	11	HSU67808
C 25	68.4	0.2	95	13	HUMUT8002B
C 26	67.8	0.2	100	9	HUMGALNSA
C 27	68	0.2	103	13	HS8IC8R
C 28	68	0.2	104	9	HUMALCE272
C 29	68	0.2	107	9	HUMALCE162
C 30	67.8	0.2	108	13	G43535
C 31	67.2	0.2	85	10	HUMHISIPR
C 32	67.4	0.2	97	9	HUMDLRA2
C 33	66.4	0.2	97	9	HUMDLRA1
C 34	66.4	0.2	97	9	HUMDLRA2
C 35	66	0.2	97	9	HUMDLRDJ
C 36	65.8	0.2	99	13	HUMUT7692A
C 37	66	0.2	100	13	G43536
C 38	66	0.2	100	13	G43538
C 39	66	0.2	110	9	HUMALCE43
C 40	65.6	0.2	80	9	HUMBRKFAE
C 41	65.6	0.2	107	11	HSU67806
C 42	65.2	0.2	79	10	S73203
C 43	65.4	0.2	110	11	HSU67807
C 44	64.6	0.2	95	10	HSSTHPKIB
C 45	64.8	0.2	96	4	NVHIS2A

ALIGNMENTS

```

RESULT 1
HUMALCE221/c HUMALCE221 103 bp ss-RNA PRI 15-APR-1994
LOCUS Human carcinoma cell-derived Alu RNA transcript, clone CE221.
DEFINITION M87896
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
    source
        1..103
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /cell_line="Ntera2D1"
        /dev_stage="embryo"
        /sex="male"
        /tissue_type="carcinoma"
BASE COUNT 25 a 27 c 33 g 18 t
ORIGIN
    Query Match 0.3%; Score 87.2; DB 9; Length 103;
    Best Local Similarity 92.0%; Pred. No. 1.1e-05;
    Matches 92; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 12183 CTGAGTGCATGGCGGAGTCTGGCTCAGACGACCTCCGCTCCGGGTTCAAGCCAT 12242
Dbb 103 CTGAGTGCATGGCGGAGTCTGGCTCAGACGACCTCCGCTCCGGGTTCAAGCCAT 44
Qy 12243 TCTCTGCTTACGCTCGGAGTACGCTGGGATTACAGGCA 12282
Dbb 43 TCTCTGCTTACGCTCGGAGTACGCTGGGATTACAGGCA 4

RESULT 2
HUMALCE162/c HUMALCE162 107 bp ss-RNA PRI 15-APR-1994
LOCUS Human carcinoma cell-derived Alu RNA transcript, clone CE162.
DEFINITION M87924
ACCESSION M87924
VERSION M87924.1 GI:174871
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 107)
AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
    source
        1..107
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /cell_line="Ntera2D1"
        /dev_stage="embryo"
        /sex="male"
        /tissue_type="carcinoma"
BASE COUNT 28 a 30 c 35 g 14 t
ORIGIN
    Query Match 0.3%; Score 87; DB 9; Length 107;
    Best Local Similarity 90.3%; Pred. No. 1.2e-05;
    Matches 93; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

RESULT 3
HSLDLRN2/c HSLDLRN2 108 bp DNA PRI 20-MAY-1992
LOCUS Human LDL-receptor gene intron 14 fragment (normal gene).
DEFINITION X05250
ACCESSION X05250
VERSION X05250.1 GI:34337
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Hovinga,J.R.,
Williamson,R. and Humphries,S.
TITLE The low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901
COMMENT See X05252 for deletion junction
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES
    source
        1..108
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        1..108
        /note="intron XIV fragment"
BASE COUNT 28 a 23 c 39 g 18 t
ORIGIN
    Query Match 0.3%; Score 87.2; DB 10; Length 108;
    Best Local Similarity 88.0%; Pred. No. 1.1e-05;
    Matches 95; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 3618 CTGGGCTCACTGCAAGCTCGCTCCGGGTTATGCAATCTCATCTCAGCCTCCAG 3677
Dbb 108 CTGGGCTCACTGCAAGCTCGCTCCGGGTTATGCAATCTCATCTCAGCCTCCAG 49
Qy 3678 AGTAGCTGGGACTACAGCGCGCCACACGCTGGCTAATTTT 3725
Dbb 48 AGTAGCTGGGACTACAGCGCGCCACACGCTGGCTAATTTT 1

RESULT 4
HSLDLRN2/c HSLDLRN2 108 bp DNA PRI 20-MAY-1992
LOCUS Human LDL-receptor gene intron 14 fragment (normal gene).
DEFINITION X05250
ACCESSION X05250
VERSION X05250.1 GI:34337
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Hovinga,J.R.,
Williamson,R. and Humphries,S.
TITLE The low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901

```

```
COMMENT See X05252 for deletion junction
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES
  Source      1..108
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              1..108
  Intron      28 a 23 c 39 g 18 t
  BASE COUNT 28 a 23 c 39 g 18 t
  ORIGIN
      0.3%; Score 83.6; DB 10; Length 108;
      Best Local Similarity 86.8%; Pred. No. 4.5e-05;
      Matches 92; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
  QY 24483 AAAAATTAGCCAGGATGGTGTGGGGCCCTATATCCAGCTAATTTGGGAGGCTGAGGC 24542
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  DB 3 AAAAATTAGCCAGGATGGTGTGGGGCCCTATATCCAGCTAATTTGGGAGGCTGAGGC 62
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  QY 24543 AGGAGAATTGCTGAACCTCGGAGGTGGAGGTTCACCTGAGGCAAG 24588
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  DB 63 AGGAGAATTGCTGAACCTCGGAGGTGGAGGTTCACCTGAGGCGAG 108
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  RESULT 5
  LOCUS HSLDLRD1 108 bp DNA PRI 20-MAY-1992
  DEFINITION Human LDL-receptor mutated gene with intron 12 deletion junction.
  ACCESSION X05249
  VERSION X05249.1 GI:34335
  KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
  SOURCE human.
  ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
  Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
  AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
  TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
  the low-density-lipoprotein-receptor gene. A possible mechanism for
  the defect in a patient with familial hypercholesterolaemia
  JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE 87161901
  COMMENT *source: hypercholesterol aemia
  See X05248 for corresponding normal gene sequence
  In the defective LDL-receptor gene the deletion occurred between two
  alu-repetitive sequences, that are in the same direction, the
  deletion eliminates exons 13 and 14 and changes the reading frame
  of the resulting spliced mRNA.
  Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
  FEATURES
    source      Location/Qualifiers
              1..108
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /cell_type="blood leukocytes from a patient with familial"
  Intron      1..108
              /note="intron XIV fragment"
  BASE COUNT 28 a 20 c 40 g 20 t
  ORIGIN
      0.3%; Score 79.8; DB 10; Length 108;
      Best Local Similarity 84.1%; Pred. No. 0.00021;
      Matches 90; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
  QY 3619 TTGGCTCACTGCAAGCTCTGCCGGTTTATGCTTCATCTCATCTCCAGCTCCAGA 3678
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  DB 107 TGCCTCACCAACCTCTGCTCTCTGGGTTCAACCATTTTCTGCTCCATCCGCTCCCA 48
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  QY 3679 GTAGCTGGGACTACAGGCGCCGCCACACGCTGGCTAATTTTTT 3725
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  DB 47 GTAGCTGGGATTACAGGCACTGCCACCACGCTGGCTAATTTTGT 1
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  RESULT 7
  LOCUS HSU67803/c 108 bp RNA PRI 01-AUG-1997
  DEFINITION Human small cytoplasmic Alu transcript.
  ACCESSION U67803
  VERSION U67803.1 GI:2289917
  KEYWORDS Alu.
  SOURCE human.
  ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
  AUTHORS Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE cDNAs derived from primary and small cytoplasmic Alu (scAlu)
  transcripts
  JOURNAL J. Mol. Biol. 271 (2), 222-234 (1997)
  MEDLINE 97415756
  REFERENCE
  AUTHORS Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE Direct Submission
  JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
```

מחזורי זמן: 100 שנה

Research Institute of Innovative Technology for the Earth 9-2
 Kizugawada Kizu-cho,
 Soraku-gun, Kyoto
 Japan, 619-02
 Phone: 07747-5-2308
 Fax: 07747-5-2321.

FEATURES

source
 1..108
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_line="HepG2"
 /clone_lib="Kiseru"
 /sex="Male"
 28 a 23 c 38 g 17 t 2 others

Query Match 0.3%; Score 74.2; DB 9; Length 108;
 Best Local Similarity 85.3%; Pred. No. 0.0019;
 Matches 93; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 12200 GATCTGGCTACAGCAACCTCGGCTCCGGGTTCAAGCCATTCTCCCTCAGCCCTC 12259
 |||||
 Db 108 GATCTGGCTACTGCAACTCTGCTCCGGGNTCAAGGACTCTCCCTCAGCCCTC 49

QY 12260 CGGAGTAGCTGGGATTACAGGATCGGCACGACACACCTGGGCTAAATTT 12308
 |||||

Db 48 CYGAGTAGCTGGGATTACA-GCATGGCCACACACNCTGGCTTTTAT 1

RESULT 11

LOCUS HSLDL112 108 bp DNA PRI 20-MAY-1992
 DEFINITION Human LDL-receptor gene intron 12 fragment (normal gene) LDL - low density lipoprotein.

ACCESSION X05248
 VERSION X05248.1 GI:34334

KEYWORDS Alu repetitive sequence; low density lipoprotein receptor; repetitive sequence.
 SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS Horsthemke, B., Beisiegel, U., Dunning, A., Hovinga, J.R.,
 Williamson, R. and Humphries, S.
 TITLE Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia

JOURNAL

MEDLINE 87161901

COMMENT Eur. J. Biochem. 164 (1), 77-81 (1987)
 see X05249 for deletion junction

DATA kindly reviewed (07-DEC-1987) by HUMPHRIES S.

FEATURES

source
 1..108
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 complement(<1..65)
 /note="Alu repeat"
 1..108
 /note="intron XII fragment"
 21 a 38 c 20 g 29 t

Query Match 0.3%; Score 73.6; DB 10; Length 108;
 Best Local Similarity 81.7%; Pred. No. 0.0024;
 Matches 85; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 12204 TTGGCTCAGCAACCTCGGCTCCGGGTTCAAGCCATTCTCCCTCAGCCCTCCGGA 12263
 |||||

Db 2 TGGCCTCACCACAACCTCTGCCTCTGGGTTCAAGCCATTCTCCCTCAGCCCTCTTA 61

QY 12264 GTAGCTGGGATTACAGGATCGGCACGACACCTCGGCTAAATTT 12307
 |||||

Db 62 GTAGCTGGGATTACAAGCATGTGCCACCACGCCCGCTGATTTT 105

RESULT 12

LOCUS HSU67807/c 110 bp RNA PRI 01-AUG-1997
 DEFINITION Human small cytoplasmic Alu transcript.
 ACCESSION U67807
 VERSION U67807.1 GI:2289921
 KEYWORDS Alu.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS Shaikh, T.H., Roy, A.M., Kim, J., Batzer, M.A. and Deininger, P.L.
 TITLE CDNAs derived from primary and small cytoplasmic Alu (sAlu) transcripts
 J. Mol. Biol. 271 (2), 222-234 (1997)

JOURNAL

MEDLINE 97415756

REFERENCE

AUTHORS Shaikh, T.H., Kim, J., Batzer, M.A. and Deininger, P.L.
 TITLE Direct Submission

JOURNAL

Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The Children's Hospital of Philadelphia, 1004F Abramson Research Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

FEATURES

source
 1..110
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="TscAlu6"
 repeat_region 1..110
 /note="sAlu"
 /rpt_family="Alu"
 /rpt_type="dispersed"
 26 a 39 c 24 g 21 t

BASE COUNT
 ORIGIN

Query Match 0.3%; Score 73.4; DB 11; Length 110;
 Best Local Similarity 83.8%; Pred. No. 0.0026;
 Matches 83; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 2864 GTAGACATGGGTTTCACTATGTGGCCAGGCTAGTTTGAACCTCTGACCTCCAGTGAT 2923
 |||||

Db 99 GGAAGATGGGTTTCACTATGTGGCCAGGCTAGTTTGAACCTCTGACCTCCAGTGAT 40

QY 2924 CCATTCTCTATTGGCTCCCAAGTCTGGGATTACAGGC 2962
 |||||

Db 39 CCACCCACTTTGGCCCTCTCANAAGTCTGGGATTACAGGC 1

RESULT 13

LOCUS HS81C8R 103 bp DNA STS 05-SEP-1991
 DEFINITION Human sequence tagged site 81C8R DNA from 19q13.
 ACCESSION X57789
 VERSION X57789.1 GI:23938
 KEYWORDS STS; myotonic dystrophy.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS 1 (bases 1 to 103)
 TITLE Direct Submission
 JOURNAL Submitted (12-FEB-1991) F.L. Aldridge, ICI Pharmaceuticals, Alderley Park, Macclesfield, Cheshire, SK10 4TG, UK
 REFERENCE 2 (bases 1 to 103)
 AUTHORS Butler, R., Riley, J.H., Ogilvie, D.J., Anand, R., Buxton, J., Davies, J., Johnson, K. and Markham, A.F.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 18:05:14 ; Search time 593.83 Seconds
(without alignments)
12218.262 Million cell updates/sec

Title: US-08-852-495C-2_COPY_1_29000
Perfect score: 29000
Sequence: 1 CACACACACACACACACA.....AACCTCGCTCTGGGTTTC 29000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 433070

Minimum DB seq length: 10

Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68.4	0.2	108	1 X12095	Human biallelic po
2	66.6	0.2	100	1 X12085	Human biallelic po
3	66.6	0.2	100	1 X12086	Human biallelic po
4	65.4	0.2	86	1 V41231	Mouse embryonic ce
5	65	0.2	100	1 X12087	Human biallelic po
6	63.4	0.2	100	1 X12087	Human biallelic po
7	63.4	0.2	100	1 X12085	Human biallelic po
8	63.4	0.2	100	1 X12086	Human biallelic po
9	62.8	0.2	100	1 Q76490	Human genome fragm
10	62.2	0.2	108	1 X12095	Human biallelic po
11	60.8	0.2	100	1 T24892	Human gene signatu
12	58.6	0.2	103	1 T20927	Human gene signatu
13	57.8	0.2	103	1 T26213	Human gene signatu
14	57.8	0.2	108	1 T26828	Human gene signatu
15	56.4	0.2	84	1 T25848	Human gene signatu
16	56.2	0.2	108	1 T25009	Human gene signatu
17	54.8	0.2	87	1 T21566	Human gene signatu
18	54.4	0.2	110	1 T25260	Human gene signatu
19	53.6	0.2	69	1 Q29016	Probe to internal
20	53	0.2	106	1 Q95210	Simple tandem repe
21	52.6	0.2	110	1 T25260	Human gene signatu
22	52.2	0.2	65	1 T25588	Human gene signatu
23	52	0.2	93	1 T25688	Human gene signatu
24	52.2	0.2	103	1 T20927	Human gene signatu
25	52	0.2	108	1 T26828	Human gene signatu
26	51.4	0.2	70	1 N60231	Normal chromosome
27	51.2	0.2	100	1 T24892	Human gene signatu
28	51.4	0.2	102	1 T20743	Human gene signatu
29	51.6	0.2	110	1 T26288	Human gene signatu
30	50.6	0.2	99	1 T23728	Human gene signatu
31	49.8	0.2	91	1 T25854	Human gene signatu
32	50	0.2	99	1 T20931	Human gene signatu
33	49.6	0.2	93	1 T22572	Human gene signatu
34	49.2	0.2	108	1 T25009	Human gene signatu

c 35	48.4	0.2	93	1 T22572	Human gene signatu
c 36	48.6	0.2	97	1 T26728	Human gene signatu
c 37	48.2	0.2	93	1 T24259	Human gene signatu
c 38	47.6	0.2	69	1 T24175	Human gene signatu
c 39	47.8	0.2	92	1 T25052	Human gene signatu
c 40	47.4	0.2	65	1 T24893	Human gene signatu
c 41	47.6	0.2	95	1 Q75099	Plasmid pOKScl8a c
c 42	47.6	0.2	109	1 T23895	Human gene signatu
c 43	47.2	0.2	69	1 Q29016	Probe to internal
c 44	47.2	0.2	85	1 T24033	Human gene signatu
c 45	47	0.2	94	1 T26403	Human gene signatu

ALIGNMENTS

RESULT 1

AC	X12095	standard; DNA; 108 BP.
DE	30-MAR-1999	(first entry)
KW	Polymorphism: biallelic; human;	forensic: paternity testing; disease;
KW	detection; phenotypic typing;	characteristic; infection; hereditary;
KW	autoimmune disease; cancer;	inflammation; drug; therapy; medicament;
OS	treatment; marker; ss.	
PN	Homo sapiens.	
PD	WO9820165-A2.	
PF	05-NOV-1997;	U20313.
PR	06-NOV-1996;	US-030455.
PA	(WHED) WHITEHEAD INST.	BIOMEDICAL RES.
PI	Hudson T, Lander ES, Wang D;	
DR	WPI; 98-286974/25.	
PT	New isolated nucleic acid segments from the human genome - used for	determining polymorphic forms for use in e.g. forensics, paternity
PT	testing or phenotypic typing for disease	
PS	Claim 1: Page 219; 310pp; English.	
CC	X10269-X12937	are human DNA fragments which contain biallelic polymorphic
CC	markers which have been isolated using the primers represented in	X09121-X10268. The base occupying the polymorphic site is indicated by
CC	the appropriate IUPAC-IUB ambiguity code. These fragments can be used in	methods for determining polymorphic forms in an individual for use in
CC	e.g. forensics, paternity testing or for phenotypic typing for diseases	such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC	muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial	hypercholesterolemia, polycystic kidney disease, hereditary
CC	spherocytosis, von Willebrand's disease, tuberculous sclerosis, hereditary	haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC	syndrome, osteogenesis imperfecta, acute intermittent porphyria,	autoimmune diseases, inflammation, cancer, diseases of the nervous
CC	system, infection by pathogenic microorganisms, and characteristics such	as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC	endurance, fertility, and susceptibility or receptivity to particular	drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC	segments can also be used to produce medicaments for the treatment or	prophylaxis of such diseases.
CC	Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;	
SQ		

Query Match 0.2%; Score 68.4; DB 1; Length 108;
Best Local Similarity 82.4%; Pred. No. 0.031;

Matches 89; Conservative 1; Mismatches 17; Indels 1; Gaps 1;

QY 2854 TGTATTTTAGTAGAGATGGGTTTCACATATGTTGGCCAGGCTAGTTTGGAACTCTCTGAC 2913

Db 1 TGTCTTTTGTAGACATGAGGTTTCTTGTGTTGGCCAGGATGGTCTCGAACTCTCTGAC 60

QY 2914 CTCACATGATCCATTCATTCATGGCCCTCCC-AAAGTCTGGGATTACAG 2960

Db 61 TTCAAGTGTCCGTCGTGCCTTGGCCCTCCCAAAAGTGTGGGATTATAG 108

```
RESULT 2
X12085/c
ID X12085 standard; DNA; 100 BP.
AC X12085;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment EST98276c.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medication;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 218; 310pp; English.
CC X12085-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, polycystic kidney disease, hereditary
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 22 A; 25 C; 22 G; 30 T;

Query Match 0.2%; Score 66.6; DB 1; Length 100;
Best Local Similarity 78.8%; Pred. No. 0.058;
Matches 78; Conservative 1; Mismatches 20; Indels 0; Gaps 0;

Qy 22975 GTGGCTCATGCTGTAATCCAGACATCTTGAGAGCTGAAGAGGAGGATCGTTGAGTC 23034
||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 99 GTGACTCACACCTATATCTCGGACCTTGGGAGCTTAGGAGGAGGATGTTTGAAC 40

Qy 23035 CGGGAGTTCAAGACATCTCGGCAACACAGCAGACCC 23073
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 39 CAGGAGCTCAAGACATCTCGGCAACACATAGCAAGATC 1

RESULT 4
X1231/c
ID V41231 standard; cDNA; 86 BP.
AC V41231;
DT 01-OCT-1998 (first entry)
DE Mouse embryonic cell EST 13-4 nucleotide sequence.
KW Embryonic stem cell; ESC; non-primate; mouse; EST; human;
KW developmental gene; transgenic animal; reporter gene; ss.
OS Mus sp.
PN W09823633-A1.
PD 04-JUN-1998.
PF 25-NOV-1997; U22335.
PR 27-NOV-1996; US-032510.
PA (CORR ) CORNELL RES FOUND INC.
PI Holtschu DL, Mark WH;
DR WPI; 98-322656/28.
PT Screening for human developmental genes - by trapping in murine
PT embryonic stem cells and analysing differential expression in vitro,
PT selecting homologous non-human primate gene and using it to isolate
PT human gene
PS Claim 37; Page 18; 60pp; English.
CC Sequences shown in V41230 to V41247 represent nucleotide sequences of
CC mouse EST from tagged cDNA clones. These are used in the method of the
CC invention of screening for human developmental genes. The method
CC comprises inserting a promoterless reporter gene into a non-primate
CC mammalian embryonic stem cell (ESC) genome and identifying cellular
CC transcripts that encode the reporter gene product. Fragments of genes
CC encoding these transcripts are cloned and sequenced. A gene encoding a
CC transcript that includes unknown sequences is selected and expression
CC level of the gene encoding the transcript, or part of it, in different
CC cell types and/or different developmental stages is detected. A gene
CC showing differential expression is selected and expression levels of a
```

CC homologous non-human primate gene, in different cell types and/or at
CC different developmental stages, using the non-primate transcript as
CC probe is detected. A homologous gene having the same pattern of
CC differential expression is selected and the non-primate gene, or part of
CC it is used to identify the homologous human gene. The ESC transcripts
CC identified by this method are used to generate transgenic animals
CC selected from rats, hamsters, rabbits, dogs, pigs, horses, cows, monkey,
CC baboon or chimpanzee for study of gene function. The method provides
CC rapid and large scale screening for human developmental genes, and
CC eliminates the need to analyse reporter gene expression in embryos.
SQ Sequence 86 BP; 16 A; 28 C; 30 G; 12 T;

Query Match 0.2%; Score 65.4; DB 1; Length 86;
Best Local Similarity 86.7%; Pred. No. 0.086; Mismatches 0; Gaps 0;
Matches 72; Conservative 0; Indels 11;

QY 9607 TGACTCTCGTCTTCTTAGGCAGACGCGCTGGATGTTAGGAGGACGCCGCTGGA 9666

DB 84 TGGCTCTCGTCTTCTTGGGCAGCAGCGGCTGGATATTGGCAGGACGCCGCTGTC 25

QY 9667 GCAATGGTCACCGGCTAGCAG 9689

DB 24 GCGATGGTCACGCGCCAGCAG 2

RESULT 5
ID X12087/C
AC X12087 standard; DNA; 100 BP.
DE Human biallelic polymorphic DNA fragment EST98276a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medication;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match 0.2%; Score 65; DB 1; Length 100;
Best Local Similarity 77.8%; Pred. No. 0.099; Mismatches 21; Indels 0; Gaps 0;
Matches 77; Conservative 1;

QY 22975 GTGGCTCATGCTTAATCCAGCACTTTGAGAGGCTGAAGAGGAGGATCGTTGAGTC 23034
DB 99 GTGACTCACACTATATYCTTGCACTTTAGGAGGCTTAGGAGGAGGATGTTTGAAC 40

QY 23035 CGGGAGTTTCAAGCATCTCTGGGCAACACAGCGAGACCC 23073

DB 39 CAGGAGCTCAGACCAKCCCTGGGAAACATAGCAAGACTC 1

RESULT 6

ID X12087 standard; DNA; 100 BP.

AC X12087;

DE Human biallelic polymorphic DNA fragment EST98276a.

KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medication;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match 0.2%; Score 63.4; DB 1; Length 100;
Best Local Similarity 76.8%; Pred. No. 0.17; Mismatches 1; Mismatches 22; Indels 0; Gaps 0;
Matches 76; Conservative 1;

QY 17671 GGTCTTACTATGTTGCCAGGCTGCTCAAACTCCTGGCTTAAGTATCTCTGTC 17730

DB 1 GAGTCTTGCTATGTTTCCAGGTTGCTTGTAGCTCTGTTTCAACAACATCTCTCTCC 60

QY 17731 TCAGCTCCCAATTTCTGGATTACTACTAGTGAGTCAC 17769

DB 61 TAAGCTCTCTAAAGTCCAGGATTATAGGTGTAGTCAC 99

RESULT 7

ID X12085 standard; DNA; 100 BP.

AC X12085;

DE Human biallelic polymorphic DNA fragment EST98276c.

KW	Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW	detection; phenotypic typing; characteristic; infection; hereditary;
KW	autoimmune disease; cancer; inflammation; drug; therapy; medication;
KS	treatment; marker; ss.
OS	Homo sapiens.
PN	WO9820165-A2.
PD	14-MAY-1998.
PF	05-NOV-1997; U20313.
PR	06-NOV-1996; US-030455.
PA	(WHED) WHITEHEAD INST BIOMEDICAL RES.
PI	Hudson T, Lander ES, Wang D;
DR	WPI; 98-286974/25.
PT	New isolated nucleic acid segments from the human genome - used for
PT	determining polymorphic forms for use in e.g. forensics, paternity
PT	testing or phenotypic typing for disease
PS	Claim 1: Page 218; 310pp; English.
CC	X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC	markers which have been isolated using the primers represented in
CC	X09121-X10268. The base occupying the polymorphic site is indicated by
CC	the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC	methods for determining polymorphic forms in an individual for use in
CC	e.g. forensics, paternity testing or for phenotypic typing for diseases
CC	such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC	mucular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC	hypercholesterolemia, polycystic kidney disease, hereditary
CC	spherocytosis, von Willebrand's disease, tuberculous sclerosis, Ehlers-Danlos
CC	haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC	syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC	autoimmune diseases, inflammation, cancer, diseases of the nervous
CC	system, infection by pathogenic microorganisms, and characteristics such
CC	as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC	endurance, fertility, and susceptibility or receptivity to particular
CC	drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC	segments can also be used to produce medicaments for the treatment or
CC	prophylaxis of such diseases.
CC	Sequence 100 BP; 22 A; 25 C; 22 G; 30 T;
SQ	
Query Match 0.2%; Score 63.4; DB 1; Length 100;	
Best Local Similarity 76.8%; Pred. No. 0.17; Indels 0; Gaps 0	
Matches 76; Conservative 1; Mismatches 22;	
QY	17671 GGCTCTACTATGTTCCCGAGGTGGTCTCAAACTCCTGGGCTTAAGTGATCCTCTGCC 17730 Db 1 GAGTC TTGCTATGTTTTCCCAGGATGCTGTGAGCTCTGTTTCAACAACATCTCTCTCC 60
QY	17731 TCAGCCTCCCAAAATTGTTGGGATTAAGTAGTGAGATCAC 17769 Db 61 TAAGCCCTCCYAAAGTCCAGGATTATAGGTGTGAGTCAC 99
RESULT 8	
X12086	
ID	X12086 standard; DNA; 100 BP.
AC	X12086;
DT	30-MAR-1999 (first entry)
DE	Human biallelic polymorphic DNA fragment EST98276b.
KW	Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW	detection; phenotypic typing; characteristic; infection; hereditary;
KW	autoimmune disease; cancer; inflammation; drug; therapy; medication;
KW	treatment; marker; ss.
OS	Homo sapiens.
PN	WO9820165-A2.
PD	14-MAY-1998.
PF	05-NOV-1997; U20313.
PR	06-NOV-1996; US-030455.
PA	(WHED) WHITEHEAD INST BIOMEDICAL RES.
PI	Hudson T, Lander ES, Wang D;
DR	WPI; 98-286974/25.
PT	New isolated nucleic acid segments from the human genome - used for
PT	determining polymorphic forms for use in e.g. forensics, paternity
PT	testing or phenotypic typing for disease
PS	Claim 1: Page 219; 310pp; English.

QY 9645 GTTAGGAGGACGC 9658
DB 14 GTTGGCAGGACGC 1

RESULT 10
X12095/c
ID X12095 standard; DNA; 108 BP.
AC X12095;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment TIGR-A003M18a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medication;
KW treatment; marker; ss.
OS Homo sapiens.
PN W0920165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, Ehlers-Danlos
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria.
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;

Query Match 0.2%; Score 62.2; DB 1; Length 108;
Best Local Similarity 79.4%; Pred No. 0.26;
Matches 85; Conservative 1; Mismatches 19; Indels 2; Gaps 1;

QY 24378 TGTATTCCAGCACTTT--GGAGGCGAGCGCGGAGATCACTTGAGGTGGGAGTTGCA 24435
DB 107 TATAATCCAGCACTTTTGGGAGCGGCAAGCGGATCACTTGAAGTCAGGAGTTGCA 48
QY 24436 GACTAGCTGGCCACATGATGAAACCCCATCTCTACTAAAAATACA 24482
DB 47 GACCATCTGTGGCCACAYAGGAACCTCATCTCTACAAAAAAGACA 1

RESULT 11
T24892/c
ID T24892 standard; cDNA to mRNA; 100 BP.
AC T24892;
DT 05-NOV-1996 (first entry)
DE Human gene signature HUMGS05998.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.

PD 01-JUN-1995.
PR 11-NOV-1994; J01916.
PA (MATS/) MATSUBARA K.
PI (OKUB/) OKUBO K.
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1720; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;

Query Match 0.2%; Score 60.8; DB 1; Length 100;
Best Local Similarity 74.7%; Pred No. 0.41;
Matches 74; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 3559 TTTTTFRTTTTGAGCGAGTCTAGCTCTGTCGCCAGCTGGAGTGGCACCATC 3618
DB 100 TTTTGTGTGTTTCAACAGAGTGTCACTCTGTCAACCCAGGCGAGTGCAATC 41
QY 3619 TTGGCTCACTGCAAGCTCTGCTCCCGGCTTTATGCCAT 3657
DB 40 TCAGCTNATTTGCAAAATCTGCTCCCAAGGTTCAAGCGAT 2

RESULT 12
T20927/c
ID T20927 standard; cDNA to mRNA; 103 BP.
AC T20927;
DT 24-JUL-1996 (first entry)
DE Human gene signature HUMGS02180.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PR 11-NOV-1994; J01916.
PA (MATS/) MATSUBARA K.
PI (OKUB/) OKUBO K.
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 758-759; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-

PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 1942: 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 84 BP; 33 A; 17 C; 15 G; 19 T;

Query Match 0.2%; Score 56.4; DB 1; Length 84;
Best Local Similarity 80.5%; Pred. No. 1.8;
Matches 66; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 17641 TAATTTTAAAGGCTTTTGTAGAGATGGGCTCTTACTATGTTGCCAGGCTGGTCTC 17700
||||||| ||| ||||||||| ||| ||||| ||| ||||||||| |||
Db 82 TAATTTTAAAGATTTTGTGTAAGACAGGGTTTCCCTATATTTCCAGGCTGGTCTG 23
||||||| ||| ||||||||| ||| ||||| ||| ||||||||| |||

QY 17701 AAATCCTGGGCTTAAGTGATC 17722
||||||| ||| ||| |||
Db 22 GAATCCTGGGCTCAAGGATC 1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: June 17, 2000, 13:46:22 ; Search time 8513.5 Seconds
(without alignments)
13806.716 Million cell updates/sec

Title: US-08-852-495C-2_COPY_1_29000
Perfect score: 29000
Sequence: 1 CACACACACACACACACACA.....AACCTGCTGCCTCCTGGGTTC 29000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:

- 1: em_est1:*
- 2: em_est2:*
- 3: em_est3:*
- 4: em_est4:*
- 5: em_est5:*
- 6: em_est6:*
- 7: em_est7:*
- 8: em_est8:*
- 9: em_est9:*
- 10: em_est10:*
- 11: em_est11:*
- 12: em_est12:*
- 13: em_est13:*
- 14: em_est14:*
- 15: em_est15:*
- 16: em_est16:*
- 17: em_est17:*
- 18: em_est18:*
- 19: em_est19:*
- 20: gb_est1:*
- 21: gb_est2:*
- 22: gb_est3:*
- 23: gb_est4:*
- 24: gb_est5:*
- 25: gb_est6:*
- 26: gb_est7:*
- 27: gb_est8:*
- 28: gb_est9:*
- 29: gb_est10:*
- 30: gb_est11:*
- 31: gb_est12:*
- 32: gb_est13:*
- 33: gb_est14:*
- 34: gb_est15:*
- 35: gb_est16:*
- 36: gb_est17:*
- 37: gb_est18:*
- 38: gb_est19:*
- 39: gb_est20:*
- 40: gb_est21:*
- 41: gb_est22:*
- 42: gb_est23:*
- 43: gb_est24:*
- 44: gb_est25:*

- 45: gb_est26:*
- 46: gb_est27:*
- 47: gb_est28:*
- 48: gb_est29:*
- 49: gb_est30:*
- 50: gb_est31:*
- 51: gb_est32:*
- 52: em_est20:*
- 53: em_est21:*
- 54: em_est22:*
- 55: em_est23:*
- 56: em_est24:*
- 57: em_est25:*
- 58: em_est26:*
- 59: gb_est33:*
- 60: gb_est34:*
- 61: gb_est35:*
- 62: gb_est36:*
- 63: gb_est37:*
- 64: gb_est38:*
- 65: em_est27:*
- 66: em_est28:*
- 67: em_est29:*
- 68: em_est30:*
- 69: gb_est39:*
- 70: gb_est40:*
- 71: gb_est41:*
- 72: gb_est42:*
- 73: gb_est43:*
- 74: gb_est44:*
- 75: em_est31:*
- 76: em_est32:*
- 77: em_est33:*
- 78: em_est34:*
- 79: gb_est45:*
- 80: gb_est46:*
- 81: gb_est47:*
- 82: gb_gss1:*
- 83: gb_gss2:*
- 84: gb_gss3:*
- 85: gb_gss4:*
- 86: em_gss1:*
- 87: em_gss2:*
- 88: em_gss3:*
- 89: em_gss4:*
- 90: gb_gss5:*
- 91: gb_gss6:*
- 92: gb_gss7:*
- 93: gb_gss8:*
- 94: gb_gss9:*
- 95: em_gss5:*
- 96: em_gss6:*
- 97: em_gss7:*
- 98: em_gss8:*
- 99: em_gss9:*
- 100: em_gss10:*
- 101: em_gss11:*
- 102: gb_gss10:*
- 103: gb_gss11:*
- 104: em_gss12:*
- 105: gb_gss12:*
- 106: gb_gss13:*
- 107: gb_gss14:*
- 108: gb_gss15:*
- 109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 727 Std Error: 0.00

Seq primer: M13RP1

High quality sequence stop: 68.

Location/Qualifiers

1. .95

/organism="Homo sapiens"

/db_xref="GDB:479462"

/db_xref="taxon:9606"

/clone="IMAGE:1217301"

/clone_lib="Soares fetal liver spleen INFLS"

/sex="male"

/dev_stage="20 week-post conception fetus"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)

with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;

1st strand cDNA was primed with a Pac I - oligo(dT) primer

[5', AACTGGAAGAAATTAATAAGATCTTTTTTTTTTTT 3'],

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Pac I and cloned into the Pac I

and Eco RI sites of the modified pT7T3 vector. Library

went through one round of normalization. Library

constructed by Bento Soares and M.Fatima Bonaldo."

31 a 23 c 26 g 12 t 3 others

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 0.3%; Score 92; DB 21; Length 95;

Matches 92; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 10297 TCTTTTAAAGTCCTGACCGCAAGTCCTGCCGCCCGCCGCAAGAGGGCTCCAGAG 10356

Db 1 TCTTTTAAAGTCCTGACCGCAAGTCCTGCCGCCCGCCGCAAGAGGGCTCCAGAG 60

QY 10357 GCAGTGACCAAGCGCAGAGAAGATGCGAAGAA 10391

Db 61 GCAGTGACCAAGCGCAGAGAAGANNCRAAGAA 95

RESULT 3

LOCUS

DEFINITION

at72g09.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone

IMAGE:2377600 3' similar to contains Alu repetitive

element; contains element MER22 repetitive element ;, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Dec 20, 1995 this sequence version replaced gi:1133644.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -400P from Gibco.

Location/Qualifiers

1. .105

FEATURES

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2377600"

/clone_lib="Barstead colon HPLRB7"

/sex="male"

/dev_stage="adult, age 25"

/lab_host="DH10B (phage resistant)"

/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a

modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer [5'

TGTTACGAATCTGAAGTGGAGCGCCGCTTTTTTTTTTTTTTTTTTTTTT

3']; double-stranded cDNA was ligated to Eco RI adaptors

[5', AATTCACTAGTAAT 3' and 5' ATTACTAGT 3'], digested

with Not I and cloned into the Not I and Eco RI sites of

the modified pT7T3 vector. Library constructed by Bob

Barstead."

17 a 35 c 27 g 26 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 0.3%; Score 92.2; DB 61; Length 105;

Matches 97; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 12154 GAGACGAGTTTCTCTCTTGTTCCTCCAGGCTGGAGTGCATCGCGCATCTTTGGCTCACA 12213

Db 1 GAGACGAGTTTCTCTCTTGTTCCTCCAGGCTGGAGTGCATCGCGCATCTTTGGCTCACC 60

QY 12214 GCAACCTCCGCTCCCGGGTTCAGCCATTCTCTCCCTCAGCCT 12258

Db 61 GCAACCTCCGCTCCCGGGTTCAGCGATTCTCTCTCCCTCAGCCT 105

RESULT 4

LOCUS

DEFINITION

AA807640

103 bp mRNA EST

05-MAR-1998

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Comment

On Jan 19, 1998 this sequence version replaced gi:2151346.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Sequencing by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 774 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 87.

Location/Qualifiers

1. .103

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1255473"

/clone_lib="NCI-CGAP_GC3"

/tissue_type="pooled germ cell tumors"

FEATURES

source

/lab_host="DH10B"
 /note="vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from 3 pooled
 germ cell tumors, and was then primed with a Not I -
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco
 RI adaptors (Pharmacia), digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pT7T3
 vector. Library is not normalized. Library was
 constructed by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 19 a 27 c 30 g 27 t
 ORIGIN

Query Match 0.3%; Score 87.6; DB 38; Length 103;
 Best Local Similarity 91.2%; Pred. No. 0.21;
 Matches 93; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 Qy 3734 AGTAGAGATGGGTTTACCGTGTAGCCAGAGCGTCTCGATCTCTTGACCTTCTGATC 3793
 Db 2 AGTAGAGATGGGTTTACCGTGTAGCCAGAGCGTCTCGATCTCTTGACCTTCTGATC 61
 Qy 3794 CGCTGGCTTGGCTCCCAAGTCTGGGATTACAGGTGTGA 3835
 Db 62 CGCTACCTCGCTCCCAAGTCTGGGATTACAGGTGTGA 103

RESULT 5
 LOCUS T77382 103 bp mRNA EST 15-MAR-1995
 DEFINITION YD72hl2.r1 Soares fetal liver spleen lNFLS Homo sapiens cDNA clone
 IMAGE:113831 5' similar to contains Alu repetitive element; mRNA
 sequence.
 T77382
 VERSION T77382.1 GI:694585
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 103)
 AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
 Holman, M., Hultman, N., Kucaba, T., Le, M., Lennon, G., Marra, M.,
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
 Trevas, E., Waterston, R., Williamson, A., Wohlmann, P. and
 Wilson, R.
 TITLE The WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT Other-ESTs: Yd72hl2.s1
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 943
 Source: IMAGE Consortium, LLNL This clone is available royalty-free
 through LLNL; contact the IMAGE Consortium (info@image.llnl.gov)
 for further information. Putative full length read
 Insert Length: 943 Std Error: 0.00
 Seq primer: M13RP1
 High quality sequence stop: 109.

FEATURES
 source
 1. .103
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="CDB:469448"
 /db_xref="taxon:9606"
 /clone="IMAGE:113831"
 /clone_lib="Soares fetal liver spleen lNFLS"
 /sex="male"
 /dev_stage="20 week-post conception fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: Liver and Spleen; vector: pT7T3D (Pharmacia)
 with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;

1st strand cDNA was primed with a Pac I - oligo(dT) primer
 [5' AACTGGAGAAATTAATAAGATCTTTTTTTTTTTTTTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 24 a 20 c 37 g 22 t
 ORIGIN

Query Match 0.3%; Score 86.6; DB 21; Length 103;
 Best Local Similarity 95.7%; Pred. No. 0.28;
 Matches 89; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 6073 TGAGAGTCTCACTCTCACTCAACCTCCCTCTCTATATTCAAGTGATCTCTTGGCTCA 6132
 Db 103 TGAGAGTCTCACTCTCACTCAACCTCCCTCTCTATATTCAAGTGATCTCTTGGCTCA 44
 Qy 6133 GCCTCCCGAGTAGCTGGGACTACAGCGGTGCAC 6165
 Db 43 GCCTCCCGAGTAGCTGGGACCACAGCGGCTTAC 11

RESULT 6
 LOCUS AA158786/c 106 bp mRNA EST 09-MAR-1998
 DEFINITION z063c11.r1 Stratagene pancreas (#937208) Homo sapiens cDNA clone
 IMAGE:591572 5' similar to contains Alu repetitive element; contains
 element P7R7 repetitive element; mRNA sequence.
 AA158786
 VERSION AA158786.1 GI:1733588
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 106)
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
 Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
 Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
 Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
 TITLE WashU-NCI human EST Project
 JOURNAL Unpublished (1997)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1406940.
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Putative full length read
 The vector to vector length is 119
 Insert Length: 926 Std Error: 0.00
 Seq primer: -28M13 rev2 from Amersham.

FEATURES
 source
 1. .106
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="GDB:4622958"
 /db_xref="taxon:9606"
 /clone="IMAGE:591572"
 /clone_lib="Stratagene pancreas (#937208)"
 /lab_host="SOLR cells (kanamycin resistant)"
 /note="Organ: pancreas; Vector: plasmid SK-; Site_1:
 EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
 Oligo dT. Pancreatic adenocarcinoma cell line. Average
 insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor
 sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5'
 CTCGAGTTTTTTTTTTTTTTT 3"

BASE COUNT 27 a 28 c 37 g 14 t
 ORIGIN

Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT
ORIGIN

34 a 27 c 32 g 17 t

Query Match 0.3%; Score 84; DB 33; Length 110;
Best Local Similarity 86.1%; Pred. No. 0.55;
Matches 93; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 12153 TGAGACGAAGTTCTCTCTTGTCCAGGCTGGAGTGCATTCAGCGCATTTGGCTCTAC 12212

Db 110 TGAATGGAGTTTGTCTCTTGTTCAGGCTGGAGTGCATTCAGCGCATTTGGCTCTAC 51

Qy 12213 AGCAACCTCCGCTCCCGGTTCAAGCATTCCTCTGCTCAGCCTCC 12260

Db 50 CACCACCTCCGCTCCCGGTTCAAGCATTCCTCTGCTCAGCCTCC 3

RESULT 14

AA835205

LOCUS AA835205 101 bp mRNA EST 23-FEB-1998
DEFINITION ak64h01.s1 Barstead pancreas HPLRB1 Homo sapiens cDNA clone
IMAGE:1412689 3', similar to contains Alu repetitive
element; contains element KER repetitive element ;, mRNA sequence.

ACCESSION

VERSION AA835205.1 GI:2908933

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 101)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project

Unpublished (1997)

On Nov 29, 1993 this sequence version replaced gi:636191.

Contact: Wilton RK

Washington University School of Medicine

444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40ml3 fwd. ET from Amersham.

FEATURES

source

1..101

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1412689"

/clone.lib="Barstead pancreas HPLRB1"

/sex="female"

/dev_stage="adult, 34 years"

/lab_host="DH10B"

/note="Organ: pancreas; Vector: pT73b-Pac (Pharmacia)
with a modified polylinker; Site.1: EcoRI; Site.2: NotI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5',

TGTTACGAATCGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[AATCGATCCTTG], digested with Not I and cloned into the

Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead."

BASE COUNT

ORIGIN

14 a 36 c 27 g 24 t

Query Match 0.3%; Score 83.4; DB 39; Length 101;
Best Local Similarity 89.1%; Pred. No. 0.67;
Matches 90; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 3569 TGAGACGGAGTCTAGCTCTGTGCGCCAGGCTGGAGTGGCAGTGCGCACCACATCTGGCTCACT 3628

Db 1 TGAGACGGAGTCTAGCTCTGTGCGCCAGGCTGGAGTGGCAGTGCGCTCACT 60

Qy 3629 GCAAGCTCTGCTCCCGGGTTTATGCCATTTCTCATGTCTCA 3669

Db 61 GCAAGCTCCGCTCCCGGGTTTACAGCCATTTCTCTGCGCTCA 101

RESULT 15

AI991750

LOCUS AI991750 106 bp mRNA EST 08-SEP-1999
DEFINITION wt48e01.x1 NCI-CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2510712 3'
similar to contains Alu repetitive element; contains element LTR8
repetitive element ;, mRNA sequence.

ACCESSION

VERSION AI991750.1 GI:5838578

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 106)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

On Dec 20, 1995 this sequence version replaced gi:1133359.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Life Technologies catalog #: 11548-013

DNA Sequencing by: Washington University Genome Sequencing Center

Clone Distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.blo.llnl.gov/bbrp/image/image.html

Seq primer: -40UP from Gibco

High quality sequence stop: 62.

Location/Qualifiers

1..106

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2510712"

/clone.lib="NCI-CGAP_Pan1"

/tissue_type="adenocarcinoma"

/lab_host="DH10B"

/note="Organ: pancreas; Vector: pCMV-SPORT6; Site.1: SalI;

Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.72 kb. Life Technologies catalog #:

11548-013"

BASE COUNT 24 a 23 c 22 g 37 t

ORIGIN

Query Match 0.3%; Score 83.6; DB 63; Length 106;

Best Local Similarity 86.8%; Pred. No. 0.62;

Matches 92; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 6383 TTGTCTAGCTTTATAGAGATGGGGTTTCGCATCTTGGCAGGCTGGCTCAAACTCCT 6442

Db 1 TTTTCTTTTAAATAGAGATGAGGTTTACCATGTTGGCAAGGTAGTCTCAAACTCCT 60

Qy 6443 GACCTCAGGTGATCTACCCACCTCAGCCTCCCAAGTCTGGGATT 6488

Db 61 GACCTCAGGTGATCTACCCACCTCAGCCTCCCAAGTCTGGGATT 106

Search completed: June 17, 2000, 20:31:38
Job time: 287665 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 17:39:10 ; Search time 372.61 seconds
(without alignments)
10116.654 Million cell updates/sec

Title: US-08-852-495C-2_COPY_1_29000
Perfect score: 29000
Sequence: 1 CACACACACACACACACA.....AACCTCTGCTCCTGGGTTTC 29000

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/6_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
7: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	69.8	0.2	105	4	US-08-481-658B-65
2	69.6	0.2	105	4	US-08-481-658B-65
3	69.8	0.2	105	4	US-08-477-504A-65
4	69.6	0.2	105	4	US-08-477-504A-65
5	69.8	0.2	105	4	US-08-486-756A-65
6	69.6	0.2	105	4	US-08-486-756A-65
7	69.8	0.2	105	4	US-08-485-862B-65
8	69.6	0.2	105	4	US-08-485-862B-65
9	69.8	0.2	105	5	US-08-787-739-65
10	69.6	0.2	105	5	US-08-787-739-65
11	58.6	0.2	78	3	US-08-454-557C-70
12	58.6	0.2	78	4	US-08-340-426D-70
13	58.6	0.2	78	4	US-08-450-673C-70
14	58.6	0.2	78	6	PCT-US95-17111A-70
15	57.2	0.2	78	3	US-08-454-557C-70
16	57.2	0.2	78	4	US-08-340-426D-70
17	57.2	0.2	78	4	US-08-450-673C-70
18	57.2	0.2	78	6	PCT-US95-17111A-70
19	54.8	0.2	76	3	US-08-454-557C-69
20	54.8	0.2	76	4	US-08-340-426D-69
21	54.8	0.2	76	4	US-08-450-673C-69
22	54.8	0.2	76	6	PCT-US95-17111A-69
23	55	0.2	84	3	US-08-454-557C-91
24	55	0.2	84	4	US-08-340-426D-91
25	55	0.2	84	4	US-08-450-673C-91
26	55	0.2	84	6	PCT-US95-17111A-91
27	54.2	0.2	84	3	US-08-454-557C-91

Sequence 91, Appl
Sequence 91, Appl
Sequence 91, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 92, Appl
Sequence 92, Appl
Sequence 92, Appl
Sequence 36, Appl
Sequence 57, Appl
Sequence 57, Appl
Sequence 57, Appl
Sequence 60, Appl
Sequence 60, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-65
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/360,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.2%; Score 69.8; DB 4; Length 105;
Best Local Similarity 79.0%; Pred. No. 9.6e-07;

CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-477-504A-65

Query Match 0.2%; Score 69.6; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. No. 1.1e-06;
Matches 89; Conservative 0; Mismatches 14; Indels 1; Gaps 1;
QY 1393 ATCTCAGCACTTTGGAGGCTGCTGCTCTACTAAATACAAAA 1451
Db 105 ATCCAGCACTTTGGAGGCCGAGGCTGTGTGATCACAAGGTCAGAGTTTGAGAGCAGC 46
QY 1452 CTGGCAATATGCGCAAAACCTGTCTCTACTAAATACAAAA 1495
Db 45 CTGGCAATATGCGCAAAACCTGTCTCTACTAAAGTGTAAAA 2

RESULT 5
US-08-486-756A-65
; Sequence 65, Application US/08486756A
; Patent No. 5981711
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,756A
; FILING DATE: 07-JUN-1995

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-486-756A-65

Query Match 0.2%; Score 69.8; DB 4; Length 105;
Best Local Similarity 79.0%; Pred. No. 9.6e-07;
Matches 83; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 3720 TTTTATTTTATTTAGTAGAGATGGGGTTTCCACCGTTAGCCAGAACGGTCTCGATCTC 3779
Db 1 TTTTATCATCTTTAGTAGAGACAGGGTTTCCACATATTTGGCCAGGCTGCTCTCAAATC 60
QY 3780 TTGACCTTCTGATCCGCCCTGCTGCTGTTCCCAAAAGTCTGGGAT 3824
Db 61 CTGACCTTGTGATCCACAGCCTCGGCCCTCCCAAAAGTCTGGGAT 105

RESULT 6
US-08-486-756A-65/c
; Sequence 65, Application US/08486756A
; Patent No. 5981711
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,756A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-486-756A-65

Query Match      0.2%; Score 69.6; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. No. 1.1e-06;
Matches 89; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 1393 ATCTCAGCACTTTGGAGGCTGAGG-GCACAGATCACGAGTCGGAGTTTGAGACCAGC 1451
Db 105 ATCCAGCACTTTGGAGGCGGAGGCTGGTGATCACAGGTCAGAGGTTTGAGAGCAGC 46

QY 1452 CTGGCAATATGGGAAACCCCTGTCTCTACTAAAAATACAAAA 1495
Db 45 CTGGCAATATGTTGAACCCCTGTCTCTACTAAAGATGTAAAA 2

RESULT 7
US-08-485-862B-65
; Sequence 65, Application US/08485862B
; Patent No. 5989838
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,862B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-485-862B-65

Query Match      0.2%; Score 69.6; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. No. 1.1e-06;
Matches 89; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 1393 ATCTCAGCACTTTGGAGGCTGAGG-GCACAGATCACGAGTCGGAGTTTGAGACCAGC 1451
Db 105 ATCCAGCACTTTGGAGGCGGAGGCTGGTGATCACAGGTCAGAGGTTTGAGAGCAGC 46

QY 1452 CTGGCAATATGGGAAACCCCTGTCTCTACTAAAAATACAAAA 1495
Db 45 CTGGCAATATGTTGAACCCCTGTCTCTACTAAAGATGTAAAA 2

RESULT 8
US-08-485-862B-65/C
; Sequence 65, Application US/08485862B
; Patent No. 5989838
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,862B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-485-862B-65

Query Match      0.2%; Score 69.6; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. No. 1.1e-06;
Matches 89; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 1393 ATCTCAGCACTTTGGAGGCTGAGG-GCACAGATCACGAGTCGGAGTTTGAGACCAGC 1451
Db 105 ATCCAGCACTTTGGAGGCGGAGGCTGGTGATCACAGGTCAGAGGTTTGAGAGCAGC 46

QY 1452 CTGGCAATATGGGAAACCCCTGTCTCTACTAAAAATACAAAA 1495
```

DB 45 CRGGCAATATGCTGAACCCGTCTCTACTAAAGATGTAAAAA 2

RESULT 9
US-08-787-739-65
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-787-739-65

Query Match 0.2%; Score 69.8; DB 5; Length 105;
Best Local Similarity 79.0%; Pred. No. 9.6e-07;
Matches 83; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 3720 TTTTATTTTATTAGTAGAGATGGGGTTTACCGTGTAGCCAGAACGGTCTCGATCTC 3779
Db 1 TTTTATTTTATTAGTAGAGATGGGGTTTACCGTGTAGCCAGAACGGTCTCGATCTC 60
QY 3780 TTGACCTTCTGATCCGCCCTTGCGTTTCCCAAAAGTGTGGGAT 3824
Db 61 CTGACCTTGTGATCCACCACCGCTCGGCCCTCCCAAAAGTGTGGGAT 105

RESULT 10
US-08-787-739-65/c
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-787-739-65

```
Query Match          0.2%; Score 69.6; DB 5; Length 105;
Best Local Similarity 85.6%; Pred. No. 1.1e-06;
Matches 89; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

Qy 1393 ATCTCAGCAGCTTTGGGAGCTGTAGG-GCACAGATCAGAGTCGGGAGTTTGACAGCAGC 1451
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 105 ATCCAGCAGCTTTGGGAGCGGAGCTGGTGCATCACAAGTCAGGAGTTTGAGAGCAGC 46

Qy 1452 CTGGCAATATATGGCGAAACCTGTCTCTACTAAATAACAAAA 1495
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 45 CTGGCAATATATGGTGAACCCCTGTCTCTACTAAAGATCTAAAAA 2

RESULT 11
US-08-454-557C-70
; Sequence 70, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-454-557C-70

Query Match          0.2%; Score 58.6; DB 3; Length 78;
Best Local Similarity 87.7%; Pred. No. 0.00025;
Matches 64; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 2840 GCCCAGCTAAATTTTGTATTTTAGTAGAGATGGGGTTTCTACTATGTTGGCCAGGCTAGT 2899
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6 GCCCAGCTAAATTTGTATTTTAGTAGAGATGGGGTTTCTCCATGTTTCATCAGGCTGTT 65

Qy 2900 TTGGAACCTCTGA 2912
      | ||||| |||||
Db 66 GTCGAACCTCCTGA 78

RESULT 12
US-08-340-426D-70
; Sequence 70, Application US/08340426D
; Patent No. 5948634
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
```

```
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-340-426D-70

Query Match          0.2%; Score 58.6; DB 4; Length 78;
Best Local Similarity 87.7%; Pred. No. 0.00025;
Matches 64; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 2840 GCCCAGCTAAATTTTGTATTTTAGTAGAGATGGGGTTTCTACTATGTTGGCCAGGCTAGT 2899
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6 GCCCAGCTAAATTTGTATTTTAGTAGAGATGGGGTTTCTCCATGTTTCATCAGGCTGTT 65

Qy 2900 TTGGAACCTCTGA 2912
      | ||||| |||||
Db 66 GTCGAACCTCCTGA 78

RESULT 13
US-08-450-673C-70
; Sequence 70, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-450-673C-70

Query Match          0.2%; Score 58.6; DB 4; Length 78;
Best Local Similarity 87.7%; Pred. No. 0.00025;
Matches 64; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2840 GCCCAGCTAATTTTGTATTTTCTAGATGGGGTTTCTACTATGTTGCCAGGCTAGT 2899
Db 6 GCCCAGCTAATTTTGTATTTTCTAGATGGGGTTTCTCTCCATGTTTCATCAGGCTGCT 65

QY 2900 TTGGAACCTCTGA 2912
Db 66 GTCGAACCTCTGA 78

RESULT 14
PCT-US95-17111A-70
; Sequence 70, Application PC/TUS951711A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-450-673C-70

Query Match          0.2%; Score 58.6; DB 4; Length 78;
Best Local Similarity 87.7%; Pred. No. 0.00025;
Matches 64; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2840 GCCCAGCTAATTTTGTATTTTCTAGATGGGGTTTCTACTATGTTGCCAGGCTAGT 2899
Db 6 GCCCAGCTAATTTTGTATTTTCTAGATGGGGTTTCTCTCCATGTTTCATCAGGCTGCT 65

QY 2900 TTGGAACCTCTGA 2912
Db 66 GTCGAACCTCTGA 78

RESULT 14
PCT-US95-17111A-70
; Sequence 70, Application PC/TUS951711A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-450-673C-70
```

```
;
; TOPOLOGY: both
; PCT-US95-17111A-70

Query Match          0.2%; Score 58.6; DB 6; Length 78;
Best Local Similarity 87.7%; Pred. No. 0.00025;
Matches 64; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2840 GCCCAGCTAATTTTGTATTTTCTAGATGGGGTTTCTACTATGTTGCCAGGCTAGT 2899
Db 6 GCCCAGCTAATTTTGTATTTTCTAGATGGGGTTTCTCTCCATGTTTCATCAGGCTGCT 65

QY 2900 TTGGAACCTCTGA 2912
Db 66 GTCGAACCTCTGA 78

RESULT 15
US-08-454-557C-70/C
; Sequence 70, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-454-557C-70

Query Match          0.2%; Score 57.2; DB 3; Length 78;
Best Local Similarity 83.3%; Pred. No. 0.00051;
Matches 65; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1433 TCGGGAGTTTGAGACCAGCCTGGCCCAATATGCGGAACCCCTGCTCTCTACTATAAATACAA 1492
Db 78 TCAGGAGTTTCGACACACGCTGATGAACATGGAGAAACCCCATCTCTACTATAAATACAA 19

QY 1493 AAATTAGCTGGCGTGT 1510
Db 18 ATATTAGCTGGCGTGT 1
```

Search completed: June 18, 2000, 01:37:37
Job time: 304987 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2000, 01:31:16 ; Search time 17971.2 Seconds
(without alignments)
-1569.844 Million cell updates/sec

Title: US-08-852-495C-2_COPY_28000_57000
Perfect score: 29001
Sequence: 1 GATAATATTATTAATTAT.....ACCTTTGATGAGAACATGA 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba1.*
- 2: gb_ba2.*
- 3: gb_om.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_p11.*
- 8: gb_p12.*
- 9: gb_p1.*
- 10: gb_p2.*
- 11: gb_p3.*
- 12: gb_ro.*
- 13: gb_sts.*
- 14: gb_sy.*
- 15: gb_un.*
- 16: gb_vi.*
- 17: em_fun.*
- 18: em_humi.*
- 19: em_hum2.*
- 20: em_in.*
- 21: em_om.*
- 22: em_or.*
- 23: em_ov.*
- 24: em_pat.*
- 25: em_ph.*
- 26: em_pl.*
- 27: em_ro.*
- 28: em_sts.*
- 29: em_sy.*
- 30: em_un.*
- 31: em_vi.*
- 32: gb_htg1.*
- 33: gb_htg2.*
- 34: gb_in1.*
- 35: gb_in2.*
- 36: em_ba1.*
- 37: em_ba2.*
- 38: em_hum3.*
- 39: em_hum4.*
- 40: gb_pr4.*
- 41: gb_htg3.*
- 42: gb_htg4.*
- 43: gb_htg5.*
- 44: gb_htg6.*

- 45: gb_htg7.*
- 46: em_htg1.*
- 47: em_htg2.*
- 48: em_htg3.*
- 49: em_hum5.*
- 50: gb_p13.*
- 51: gb_pr5.*
- 52: gb_htg8.*
- 53: gb_htg9.*
- 54: gb_htg10.*
- 55: gb_htg11.*
- 56: gb_htg12.*
- 57: gb_htg13.*
- 58: gb_htg14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	94.4	0.3	96	13	G31304	G31304 sy899g1-19
2	91.6	0.3	108	10	HSJDLRN2	X05250 Human LDL-r
3	88.8	0.3	108	10	HSJDLRN2	X05250 Human LDL-r
4	88	0.3	108	11	HSU67803	U67803 Human small
5	87	0.3	107	9	HUMALCE162	M87924 Human carci
6	85.4	0.3	108	11	HSU67804	U67804 Human small
7	85.4	0.3	103	9	HUMALCE221	M87896 Human carci
8	84	0.3	103	9	HUMALCE221	M87896 Human carci
9	84.2	0.3	108	10	HSJDLRD1	X05249 Human LDL-r
10	84.2	0.3	108	10	HSJDLRD2	X05251 Human LDL-r
11	81.4	0.3	108	10	HSJDLI12	X05248 Human LDL-r
12	81.4	0.3	108	10	HSJDLRD1	X05249 Human LDL-r
13	81.4	0.3	108	10	HSJDLRD2	X05251 Human LDL-r
14	79	0.3	108	11	HSU67808	U67808 Human small
15	78.8	0.3	110	9	HUMALCE43	M87900 Human carci
16	78.2	0.3	110	11	HSU67807	U67807 Human small
17	77.4	0.3	90	9	HUMDLRFL	X03555 Human low d
18	77.6	0.3	103	13	HS8IC8R	X57789 Human sequ
19	77.6	0.3	104	9	HUMALCE272	M87899 Human carci
20	77.4	0.3	107	9	HUMALCE162	M87924 Human carci
21	76.2	0.3	108	11	HSU67803	U67803 Human small
22	75.4	0.3	97	9	HUMDLRD3	M14179 Human famil
23	75.6	0.3	103	13	HS8IC8R	X57789 Human sequ
24	75.6	0.3	108	10	HSJDLI12	X05248 Human LDL-r
25	75	0.3	110	11	HSU67807	U67807 Human small
26	73.8	0.3	97	9	HUMDLRA1	M14178 Human low d
27	73.6	0.3	97	9	HUMDLRA2	M14180 Human low d
28	73	0.3	107	11	HSU67806	U67806 Human small
29	73.2	0.3	110	9	HUMALCE43	M87900 Human carci
30	72.6	0.3	100	9	HUMGALNSA	D45223 Human GALNS
31	72.6	0.3	108	9	HUMD1D03M5	D16965 Human HepG2
32	72.2	0.2	108	9	HUMD1D03M5	D16965 Human HepG2
33	71	0.2	104	9	HUMALCE272	M87899 Human carci
34	71	0.2	107	11	HSU67806	U67806 Human small
35	70.6	0.2	101	10	S79560	S79560 HRX {intron
36	69.8	0.2	108	11	HSU67808	U67808 Human small
37	69.2	0.2	91	13	HUMUT8164A	L30244 Human STS U
38	69.2	0.2	99	13	HUMUT7692A	L30306 Human STS U
39	69.2	0.2	100	13	HUMUT931A	L31299 Human STS U
40	69.4	0.2	108	13	G43535	G43535 WIAP-2393-S
41	68.8	0.2	80	9	HUMERKFAE	M36135 Human alpha
42	68.8	0.2	108	13	G32614	G32614 A009K21 Hum
43	67.6	0.2	91	13	HUMUT8164A	L30244 Human STS U
44	67.8	0.2	100	9	HUMGALNSA	D45223 Human GALNS
45	67.6	0.2	107	13	G32919	G32919 A009W27 Hum

ALIGNMENTS

```

RESULT 1
G31304
LOCUS G31304 96 bp DNA STS 29-SEP-1998
DEFINITION sy899g1-19 Human (A.Gnirke) Homo sapiens STS genomic, sequence
tagged site.
ACCESSION G31304
VERSION G31304.1 GI:1871333
KEYWORDS STS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 96)
AUTHORS Lauer,P., Meyer,N.C., Prass,C.E., Starnes,S.M., Wolff,R.K. and
Gnirke,A.
TITLE Clone-contig and STS maps of the hereditary hemochromatosis region
on human chromosome 6p21.3-6p22
JOURNAL on human chromosome 6p21.3-6p22
MEDLINE Genome Res. 7 (5), 457-470 (1997)
COMMENT 97294058

GDB: GDB:5584195
GDB_DSEG: D6S2377
Contact: Andreas Gnirke
Mercator Genetics, Inc.
4040 Campbell Ave, Menlo Park CA, 94025, USA
Email: gnirke@mercator.com
Primer A: GTCCCAAGAAATATAAATCAG
Primer B: AGGCACAGTGGGAAG
STS size: 77
PCR Profile:
Denaturation: 92 degrees C for 20 seconds
Annealing: 60 degrees C for 45 seconds
Polymerization: 72 degrees C for 60 seconds
PCR Cycles: 35
Thermal Cycler: MJ Research PTC-200

Protocol:
Template: 30-200 ng
Primer: each 0.8 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 12 ul

Buffer:
MgCl2: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3
gelatin: 0.001% (w/v).
Location/Qualifiers
1..96
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human (A.Gnirke)"
9..85
primer_bind 9..30
primer_bind complement(70..85) 22 t
BASE COUNT 22 a 29 c 23 g 22 t
ORIGIN

STS
primer_bind 9..30
primer_bind complement(70..85) 22 t
BASE COUNT 22 a 29 c 23 g 22 t
ORIGIN

Query Match 0.3%; Score 94.4; DB 13; Length 96;
Best Local Similarity 99.0%; Pred. No. 4.2e-08;
Matches 95; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12823 CTGCTGGTGTCCCAAGAAATATAAATGAGAAAATGCTTCCATGGATGCCAGATCCCC 12882
|||||
Db 1 CTGCTGGTGTCCCAAGAAATATAAATGAGAAAATGCTTCCATGGATGCCAGATCCCC 60

QY 12883 TCTGCCCTCTTCCACATGTGCTGGGGGAGAGGT 12918
|||||
Db 61 TCTGCCCTCTTCCACATGTGCTGGGGGAGAGGT 96

```

```

RESULT 2
HSLDLRN2/c
LOCUS HSLDLRN2 108 bp DNA PRI 20-MAY-1992
DEFINITION Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION X05250
VERSION X05250.1 GI:34337
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
Williamson,R., and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901
COMMENT See X05252 for deletion junction
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES
source
Location/Qualifiers
1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..108
/note="intron XIV fragment"
BASE COUNT 28 a 23 c 39 g 18 t
ORIGIN

Query Match 0.3%; Score 91.6; DB 10; Length 108;
Best Local Similarity 91.5%; Pred. No. 1.4e-07;
Matches 97; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 10742 CTCGGCTCACTGCAAGCTCTGCTCTCTGGTTTCATGCCATTTCTCGCTCAGCCTCCG 10801
|||||
Db 108 CTCGGCTCACTGCAAGCTCTGCTCTCTGGTTTCATGCCATTTCTCGCTCAGCCTCCG 49

QY 10802 AGTAGCTGGGACTACAGCGTCTGCCACACGCCAGCTAATTTTT 10847
|||||
Db 48 AGTAGCTGGGATTACAGGCACCTGCCACACGCCCTGGCTAATTTTT 3

RESULT 3
HSLDLRN2
LOCUS HSLDLRN2 108 bp DNA PRI 20-MAY-1992
DEFINITION Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION X05250
VERSION X05250.1 GI:34337
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
Williamson,R., and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901
COMMENT See X05252 for deletion junction
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES
source
Location/Qualifiers
1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..108
/note="intron XIV fragment"
BASE COUNT 28 a 23 c 39 g 18 t
ORIGIN

```


ORIGIN

Query Match 0.3%; Score 88.8; DB 10; Length 108;
Best Local Similarity 88.9%; Pred. No. 4.7e-07;
Matches 96; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 7126 AAAAACTTAGCGTCATGGTGGCCCTGTAGTCTTCAGCCACTTGGGAGGCTGAG 7185

Db 1 ACAAATATGACGCGTGGTGGCAGGTCCCTGTAATCCCACTACTCGGAGGCTGAG 60

QY 7186 GCAGGAAATGCTTGAACCCAGGAGGAGGTTGCAGTGAGCCCGAG 7233

Db 61 GCAGGAAATGCTTGAACCCAGGAGGAGGTTGCAGTGAGCCCGAG 108

RESULT 4

HSU67803/c 108 bp RNA PRI 01-AUG-1997

LOCUS Human small cytoplasmic Alu transcript.

DEFINITION U67803

ACCESSION U67803.1 GI:2289917

VERSION Alu.

KEYWORDS human.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 108)

AUTHORS Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.

TITLE cDNAs derived from primary and small cytoplasmic Alu (scAlu)

transcripts

J. Mol. Biol. 271 (2), 222-234 (1997)

MEDLINE 97415756

AUTHORS Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.

TITLE Direct Submission

Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The

Children's Hospital of Philadelphia, 1004F Abramson Research

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA


```

VERSION X05251.1 GI:34336
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
          Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
          Williamson,R. and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
        the low-density-lipoprotein-receptor gene. A possible mechanism for
        the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901
COMMENT *source: hypercholesterol aemia
        See X05250 for corresponding normal gene sequence
        In the defective LDL-receptor gene the deletion occurred between two
        alu-repetitive sequences, that are in the same direction, the
        deletion eliminates exons 13 and 14 and changes the reading frame
        of the resulting spliced mRNA.
        Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES
    source      Location/Qualifiers
    1..108
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /cell_type="blood leukocytes from a patient with familial"
    Intron
    1..108
    /note="intron XIV fragment"
    BASE COUNT 28 a 20 c 40 g 20 t
    ORIGIN

Query Match      0.3%; Score 84.2; DB 10; Length 108;
Best Local Similarity 87.6%; Pred. No. 3 5e-06;
Matches 92; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 10743 TCGGCTCACTGCAAGCTCTCCCTCCCTGGGTTCATGCAATTCCTCGCTCAGCCTCCCGA 10802
      ||| |||| ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 107 TCGCCTCACCACACACCTCTCCCTCCCTGGGTTCACAAACCAATTTCTCGCTCAGCCTCCCGA 48

QY 10803 GTAGCTGGGATTACAGCGTCTGCCACCCAGCCCGCAGCTAATTTT 10847
      ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 47 GTAGCTGGGATTACAGCGACCTGCCACCCAGCGCTGGCTAATTTT 3

RESULT 11
HSLDLI12
LOCUS Human LDL-receptor gene intron 12 fragment (normal gene) LDL = low
        density lipoprotein.
DEFINITION X05248
ACCESSION X05248
VERSION X05248.1 GI:34334
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor;
        repetitive sequence.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
          Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
          Williamson,R. and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
        the low-density-lipoprotein-receptor gene. A possible mechanism for
        the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901
COMMENT see X05249 for deletion junction
        Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES
    source      Location/Qualifiers
    1..108
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    complement(<1..65)

```

```

Intron
1..108
/note="Alu repeat"
BASE COUNT 21 a 38 c 20 g 29 t
ORIGIN

Query Match      0.3%; Score 81.4; DB 10; Length 108;
Best Local Similarity 85.0%; Pred. No. 1.2e-05;
Matches 91; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 5414 TCGGCTCACCAGCAACCTCTACCTCCAGGTTCAAGCAATTCCTCGCTCAGCCTCCCGA 5473
      ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 2 TCGCCTCACCACCAACCTCTCGCTCCCTGGGTTCAAACCAATTTCTCGCTCAGCCTCCTTA 61

QY 5474 GTAGCTGGGATTACAGCATGCATCACCAGCCCGCAGCTAATTTTGA 5520
      ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 62 GTAGCTGGGATTACAGCATGTGCCACCCAGCGCTGATTTTGA 108

RESULT 12
HSLDLRD1/c
LOCUS Human LDL-receptor mutated gene with intron 12 deletion junction.
DEFINITION X05249
ACCESSION X05249
VERSION X05249.1 GI:34335
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
          Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
          Williamson,R. and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
        the low-density-lipoprotein-receptor gene. A possible mechanism for
        the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901
COMMENT *source: hypercholesterol aemia
        See X05248 for corresponding normal gene sequence
        In the defective LDL-receptor gene the deletion occurred between two
        alu-repetitive sequences, that are in the same direction, the
        deletion eliminates exons 13 and 14 and changes the reading frame
        of the resulting spliced mRNA.
        Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES
    source      Location/Qualifiers
    1..108
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /cell_type="blood leukocytes from a patient with familial"
    misc_feature 1..108
    /note="deletion junction region intron 12/ intron 15"
    BASE COUNT 20 a 40 c 20 g 28 t
    ORIGIN

Query Match      0.3%; Score 81.4; DB 10; Length 108;
Best Local Similarity 85.0%; Pred. No. 1.2e-05;
Matches 91; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 7126 AAAAATTTAGCGCTGCTGGCATGCCCTGTACTCTCAGCCACTTGGGAGGCTGAG 7185
      ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 108 ACAAAATTTAGCAGCGCTGCTGGCAGGTCCCTGTATATCCAGCTACTCCGGAGGCTGAG 49

QY 7186 GCAGGAAATTTGCTTGAACCCAGGAGGAGGTTGCAGTGAGCCGA 7232
      ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 48 GCAGGAAATTTGTTGAACCCAGGAGGAGGTTGTGGTGAGGCCGA 2

RESULT 13
HSLDLRD2
LOCUS

```

DEFINITION Human LDL-receptor mutated gene with intron 14 deletion junction.
ACCESSION X05251
VERSION X05251.1 GI:34336
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
AUTHORS Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 108)
TITLE Horsthemke, B., Beisiegel, U., Dunning, A., Havinga, J.R.,
Williamson, R., and Humphries, S.
Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia
Eur. J. Biochem. 164 (1), 77-81 (1987)
JOURNAL 87161901
MEDLINE
COMMENT *source: hypercholesterol aemia
See X05250 for corresponding normal gene sequence
In the defective LDL-receptor gene the deletion occurred between two
alu-repetitive sequences, that are in the same direction, the
deletion eliminates exons 13 and 14 and changes the reading frame
of the resulting spliced mRNA.
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES
Location/Qualifiers
1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="blood leukocytes from a patient with familial"
1..108
/note="intron XIV fragment"
BASE COUNT 28 a 20 c 40 g 20 t
ORIGIN
Query Match 0.3%; Score 81.4; DB 10; Length 108;
Best Local Similarity 85.0%; Pred. No. 1.2e-05;
Matches 91; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Qy 7126 AAAAACTTAGCGGTGATGGGATGGGCTGTGCTCAGCCACTGGGAGGCTGAG 7185
Db 1 ACAAATATTAGCAGCGGTGGGAGGTGCTGTAAATCCAGCTACTCGGGAGGCTGAG 60
Qy 7186 GCAGGAAATGCTTGAACCCAGGAGGAGGTGTCAGTACGAGCCGA 7232
Db 61 GCAGGAAATGCTTGAACCCAGGAGGAGGTGTCAGTACGAGCCGA 107
RESULT 14
HSU67808/c
LOCUS HSU67808 108 bp RNA PRI 01-AUG-1997
DEFINITION Human small cytoplasmic Alu transcript.
ACCESSION U67808
VERSION U67808.1 GI:2289922
KEYWORDS Alu.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 108)
TITLE Shaikh, T.H., Roy, A.M., Kim, J., Batzer, M.A. and Deininger, P.L.
cDNAs derived from primary and small cytoplasmic Alu (scAlu)
transcripts
J. Mol. Biol. 271 (2), 222-234 (1997)
JOURNAL 97415756
MEDLINE
REFERENCE 2 (bases 1 to 108)
AUTHORS Shaikh, T.H., Kim, J., Batzer, M.A. and Deininger, P.L.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
Children's Hospital of Philadelphia, 1004F Abramson Research
Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES
Location/Qualifiers
1..108
/organism="Homo sapiens"

/db_xref="taxon:9606"
repeat_region 1..108
/clone="TscAlu7"
/note="scAlu"
/rpt_family="Alu"
/rpt_type="dispersed"
BASE COUNT 22 a 28 g 21 t
ORIGIN
Query Match 0.3%; Score 79; DB 11; Length 108;
Best Local Similarity 89.5%; Pred. No. 3.4e-05;
Matches 85; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 10860 AGAGATGGGTTCACCATGTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCAC 10919
Db 95 AAAGACGGAGTTTCACCATGTGGCCAGGCTGGTCTCAAACTCTTGACCTTGTGTATCCAC 36
Qy 10920 CCGCTTTGGCCTCCCAAGTGTGGGATTACAGGC 10954
Db 35 CCGACTTGGCTCCCAAGTGTGGGATTACAGGC 1
RESULT 15
HUMALCE43/c
LOCUS HUMALCE43 110 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE43.
ACCESSION M87900
VERSION M87900.1 GI:174876
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 110)
TITLE Sinnett, D., Richer, C., Deragon, J.-M. and Labuda, D.
Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
J. Mol. Biol. (1992) In press
FEATURES Location/Qualifiers
1..110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="NTERA2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT 27 a 31 c 34 g 18 t
ORIGIN
Query Match 0.3%; Score 78.8; DB 9; Length 110;
Best Local Similarity 84.0%; Pred. No. 3.7e-05;
Matches 89; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
Qy 10869 GTTTCACCATGTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCACCCGCTTGG 10928
Db 110 GTTTCGTCATGTAGCCAGGCTGGTCTTGAATCTACTAGCTGCAATCCTCCTGCTTGG 51
Qy 10929 CTTCCCAAGTGTGGGATTACAGGCGGTGAGCCACCGTGCCGGCC 10974
Db 50 CTTCCCAAGTGTGGGATTGTAGGTGTGAGCCACCGCCGGCC 5
Search completed: June 18, 2000, 09:50:49
Job time: 335533 sec

Copyright (c) 1993 - 2000 CompuGen Ltd.

(without alignments)
12231.247 Million cell updates/sec

Sequence: 1 GATAATTATTTTAATATTAT.....ACCCCTTTGATGAGAACATGA 29001

Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 433070

Maximum DB seq length: 110

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
C 1	70	0.2	108	1	X12095	Human biallelic po	
C 2	67.2	0.2	100	1	T24892	Human gene signatu	
C 3	66	0.2	103	1	T20927	Human gene signatu	
C 4	65.2	0.2	108	1	X12086	Human biallelic po	
C 5	64	0.2	100	1	X12086	Human biallelic po	
C 6	63.6	0.2	100	1	X12087	Human biallelic po	
C 7	63.6	0.2	100	1	X12085	Human biallelic po	
C 8	62.4	0.2	100	1	T24892	Human gene signatu	
C 9	62.6	0.2	103	1	T26213	Human gene signatu	
C 10	62.2	0.2	108	1	T26828	Human gene signatu	
C 11	61	0.2	99	1	T20931	Human gene signatu	
C 12	60.2	0.2	103	1	T20927	Human gene signatu	
C 13	59.4	0.2	91	1	T25854	Human gene signatu	
C 14	59.4	0.2	108	1	T26828	Human gene signatu	
C 15	58.6	0.2	108	1	T25009	Human gene signatu	
C 16	57.8	0.2	108	1	T25009	Human gene signatu	
C 17	57.2	0.2	91	1	T25854	Human gene signatu	
C 18	55.6	0.2	103	1	T26213	Human gene signatu	
C 19	55.2	0.2	87	1	T21566	Human gene signatu	
C 20	55.2	0.2	93	1	T25688	Human gene signatu	
C 21	55.4	0.2	110	1	T23260	Human gene signatu	
C 22	55	0.2	93	1	T23572	Human gene signatu	
C 23	53.4	0.2	87	1	T21566	Human gene signatu	
C 24	53.2	0.2	93	1	T23572	Human gene signatu	
C 25	53.4	0.2	97	1	T26728	Human gene signatu	
C 26	53	0.2	95	1	T23131	Human gene signatu	
C 27	53	0.2	109	1	T23895	Human gene signatu	
C 28	52.4	0.2	70	1	N60231	Normal chromosome	
C 29	52.6	0.2	93	1	T24259	Human gene signatu	
C 30	52.4	0.2	110	1	T25260	Human gene signatu	
C 31	52	0.2	69	1	Q29016	Probe to internal	
C 32	52	0.2	84	1	T25848	Human gene signatu	
C 33	52	0.2	93	1	T25688	Human gene signatu	
C 34	52.2	0.2	100	1	X12087	Human biallelic po	

```
RESULT 2
T24892/c
ID T24892 standard; cDNA to mRNA; 100 BP.
AC T24892;
DE 05-NOV-1996 (first entry)
DE Human gene signature HUMG506998.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1720; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;

Query Match 0.28; Score 67.2; DB 1; Length 100;
Best Local Similarity 78.8%; Pred. No. 0.023;
Matches 78; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 15731 TTTTGTCTACTGACAGCTCACTATCACCAGGCTGGAGTGCAGTGGCACAATC 15790
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 100 TTTGTTTCTTCAACACAGTGTCACTCTGTCACCCAGGCGGAGTGCANGGTGCAATC 41

Qy 15791 TCAGTCTACTGCAACCTGCACCTCCTCTGGGTTCAAGGGAT 15829
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 40 TCAGCTNATTGCAAAATCTGCTCCAGCGTTCAAGCGAT 2

RESULT 3
T20927/c
ID T20927 standard; cDNA to mRNA; 103 BP.
AC T20927;
DE 24-JUL-1996 (first entry)
DE Human gene signature HUMG502180.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1720; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;

Query Match 0.28; Score 67.2; DB 1; Length 100;
Best Local Similarity 78.8%; Pred. No. 0.023;
Matches 78; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 15731 TTTTGTCTACTGACAGCTCACTATCACCAGGCTGGAGTGCAGTGGCACAATC 15790
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 100 TTTGTTTCTTCAACACAGTGTCACTCTGTCACCCAGGCGGAGTGCANGGTGCAATC 41

Qy 15791 TCAGTCTACTGCAACCTGCACCTCCTCTGGGTTCAAGGGAT 15829
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 40 TCAGCTNATTGCAAAATCTGCTCCAGCGTTCAAGCGAT 2

RESULT 4
X12095
ID X12095 standard; DNA; 108 BP.
AC X12095;
DE 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment TIGR-A003M18a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; heredity;
KW autoimmune disease; cancer; inflammation; drug; therapy; medication;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
```



```
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 103 BP; 22 A; 27 C; 21 G; 31 T;

Query Match 0.2%; Score 60.2; DB 1; Length 103;
Best Local Similarity 74.7%; Pred. No. 0.28;
Matches 74; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 21738 APTCTCCCTGCTCCAGCTCCCAAGCAGCTGGGATTACAGGTACCTGCCACCATGCTGGT 21797
|| |||| | | |||| |||| | | |||| |||| | | |||| |||| | |
Dy 2 APTCTCCCACTCCACCTCCCAAGTAGCTGTGGCTACAGGTGTGGCCACCATGTCCAGC 61
|| |||| | | |||| |||| | | |||| |||| | | |||| |||| | |

QY 21798 TAATTTTGTATTATTAGTAGAGACGGGGTTTCACCATG 21836
| |||| |||| |||| |||| | | |||| |||| | | |||| |||| | |
Dy 62 TGATTTTGTATTATTAGTAGGACAGTATTCTCCATG 100
|| |||| |||| |||| |||| | | |||| |||| | | |||| |||| | |

RESULT 13
T25854/c
ID T25854 standard; cDNA to mRNA; 91 BP.
AC T25854;
DE 22-OCT-1996 (first entry)
DE Human gene signature HUMGS08084.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1944; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 91 BP; 18 A; 22 C; 28 G; 18 T;

Query Match 0.2%; Score 59.4; DB 1; Length 91;
Best Local Similarity 77.5%; Pred. No. 0.37;
Matches 69; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 15741 TGAGACAGTCTCTACTATATACCCAGCTGGAGTGGCGACATCTCAGTCTACT 15800
|||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Dy 90 TGAGACAGTCTCTACTATATACCCAGCTGGAGTGGCGACATCTCAGTCTACT 31
|||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

QY 15801 GCAACCTGCACCTCTCTGGTTCAAGGAT 15829
|||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
```

```
Dy 30 TGAACCCNCTGCCTCCTAGGCTCAAGTGAT 2
RESULT 14
T26828/c
ID T26828 standard; cDNA to mRNA; 108 BP.
AC T26828;
DE 14-NOV-1996 (first entry)
DE Human gene signature HUMGS09078.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 2182; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 18 A; 33 C; 23 G; 28 T;

Query Match 0.2%; Score 59.4; DB 1; Length 108;
Best Local Similarity 83.5%; Pred. No. 0.37;
Matches 66; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 6071 CAGCCGGTGCAGTGGCTCATGACTGTATCCAGCACTTTGGGAGGTCGAGGCGGTG 6130
|||||| | |||| |||| | |||| | |||| | |||| | |||| | |||| |
Dy 79 CAGCCGGGCGTGGTGGCTCATGCTGTAAACCCAGCACTATGGGAGCGGANACGGGGG 20
|||||| | |||| |||| | |||| | |||| | |||| | |||| | |||| |

QY 6131 ATCAGGAGTCCAGGAGTTC 6149
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Dy 19 ATGACGAGGTGAGGAGATC 1

RESULT 15
T25009/c
ID T25009 standard; cDNA to mRNA; 108 BP.
AC T25009;
DE 07-NOV-1996 (first entry)
DE Human gene signature HUMGS07131.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
```

PI	Matsubara K., Okubo K;
DR	WRI; 95-206931/27.
PT	Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT	for diagnosis of abnormal cell function, by preparing cDNA that
PT	reflects relative abundance of corresp. mRNA in specific human
PT	tissues
PS	Claim 1; Page 1748; 2245pp; Japanese.
CC	A single-stranded DNA (or its complementary strand or the corresp.
CC	double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC	given in T19001-T26837 and which is able to hybridise to part of
CC	human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC	sequences were obtained from 3'-directed cDNA libraries prepared
CC	from various human tissues; synthesis of cDNA was initiated from the
CC	3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC	untranslated sequence is unique to a particular mRNA species, almost
CC	all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC	is constructed so as to reflect accurately the relative abundance of
CC	different mRNAs in the particular tissue from which it was derived.
CC	The appearance frequency of a given GS in a cDNA library can be
CC	determined (esp. using primers and probes derived from the GS
CC	sequences) as a means of diagnosing abnormal cell function or for
CC	recognising different cell types.
SQ	Sequence 108 Bp; 34 A; 26 G; 15 T;
	Query Match 0.2%; Score 58.6; DB 1; Length 108;
	Best Local Similarity 71.0%; Pred. No. 0.49;
	Matches 76; Conservative 0; Mismatches 31; Indels 0; Gaps
QY	15723 TTTTTTTTTTTTTTTTTGAGACAGACTCTATCATACCAGGCTGGAGGTGCAGTG 15782
DB	108 TTGTGTTGTGTTGTTTCACAGGCTTGCTCTCACTCAGCGTGAARNACGTG 49
QY	15783 GCACAATCTCAGCTCAGTCAGCAACCTGCACCTCTCTGGGTTCAAGGAT 15929
Db	48 CGCTGACCATTGGCTCAGTCGACCGCTTGGCCCTCATGGGCTCAGGCCAT 2

Search completed: June 18, 2000, 10:13:39
Job time: 335730 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 20:31:38 ; Search time 8513.82 Seconds
(without alignments)
13806.673 Million cell updates/sec

Title: US-08-852-495C-2_COPY_28000_57000
Perfect score: 29001
Sequence: 1 GATATTATTATTAATTAT.....ACCTTTGATGAGAACATGA 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues
Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :	EST:*
	1: em_est1:*
	2: em_est2:*
	3: em_est3:*
	4: em_est4:*
	5: em_est5:*
	6: em_est6:*
	7: em_est7:*
	8: em_est8:*
	9: em_est9:*
	10: em_est10:*
	11: em_est11:*
	12: em_est12:*
	13: em_est13:*
	14: em_est14:*
	15: em_est15:*
	16: em_est16:*
	17: em_est17:*
	18: em_est18:*
	19: em_est19:*
	20: gb_est1:*
	21: gb_est2:*
	22: gb_est3:*
	23: gb_est4:*
	24: gb_est5:*
	25: gb_est6:*
	26: gb_est7:*
	27: gb_est8:*
	28: gb_est9:*
	29: gb_est10:*
	30: gb_est11:*
	31: gb_est12:*
	32: gb_est13:*
	33: gb_est14:*
	34: gb_est15:*
	35: gb_est16:*
	36: gb_est17:*
	37: gb_est18:*
	38: gb_est19:*
	39: gb_est20:*
	40: gb_est21:*
	41: gb_est22:*
	42: gb_est23:*
	43: gb_est24:*
	44: gb_est25:*
	45: gb_est26:*
	46: gb_est27:*
	47: gb_est28:*
	48: gb_est29:*
	49: gb_est30:*
	50: gb_est31:*
	51: gb_est32:*
	52: em_est20:*
	53: em_est21:*
	54: em_est22:*
	55: em_est23:*
	56: em_est24:*
	57: em_est25:*
	58: em_est26:*
	59: gb_est33:*
	60: gb_est34:*
	61: gb_est35:*
	62: gb_est36:*
	63: gb_est37:*
	64: gb_est38:*
	65: em_est27:*
	66: em_est28:*
	67: em_est29:*
	68: em_est30:*
	69: gb_est39:*
	70: gb_est40:*
	71: gb_est41:*
	72: gb_est42:*
	73: gb_est43:*
	74: gb_est44:*
	75: em_est31:*
	76: em_est32:*
	77: em_est33:*
	78: em_est34:*
	79: gb_est45:*
	80: gb_est46:*
	81: gb_est47:*
	82: gb_gss1:*
	83: gb_gss2:*
	84: gb_gss3:*
	85: gb_gss4:*
	86: em_gss1:*
	87: em_gss2:*
	88: em_gss3:*
	89: em_gss4:*
	90: gb_gss5:*
	91: gb_gss6:*
	92: gb_gss7:*
	93: gb_gss8:*
	94: gb_gss9:*
	95: em_gss5:*
	96: em_gss6:*
	97: em_gss7:*
	98: em_gss8:*
	99: em_gss9:*
	100: em_gss10:*
	101: em_gss11:*
	102: gb_gss10:*
	103: gb_gss11:*
	104: em_gss12:*
	105: gb_gss12:*
	106: gb_gss13:*
	107: gb_gss14:*
	108: gb_gss15:*
	109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result % Query

No.	Score	Match	Length	DB	ID	Description
1	93	0.3	109	30	AA243009	AA243009 zr25h02.s
2	93	0.3	109	84	B17434	B17434 345K2.TVB C
3	92.2	0.3	105	61	A1832832	A1832832 at72909.x
4	91.6	0.3	106	37	AA703692	AA703692 ag81a10.f
5	91.6	0.3	106	105	AQ264176	AQ264176 CITBI-EI-
6	91.4	0.3	109	94	AQ028426	AQ028426 CIT-HSP-2
7	90.2	0.3	103	108	AQ535244	AQ535244 RPII-11-3
8	89.4	0.3	107	35	AA565533	AA565533 nk42b11.s
9	89	0.3	105	28	AA078003	AA078003 7H12008 C
10	89	0.3	105	61	A1832832	A1832832 at72909.x
11	88.6	0.3	103	84	B48914	B48914 RPII-11-4
12	88.6	0.3	103	108	AQ535244	AQ535244 RPII-11-3
13	88.2	0.3	109	105	AQ265749	AQ265749 CITBI-EI-
14	87.6	0.3	102	94	AQ004934	AQ004934 CIT-HSP-2
15	87.6	0.3	103	38	AA807640	AA807640 nx08b05.s
16	87.6	0.3	110	64	AA083640	AA083640 xc49f02.x
17	87.2	0.3	110	33	AA442529	AA442529 zv68b02.f
18	86.6	0.3	102	30	AA226656	AA226656 nc19f09.s
19	86.8	0.3	106	94	AQ062963	AQ062963 CIT-HSP-2
20	86.8	0.3	110	39	AA897366	AA897366 am06b02.s
21	86.6	0.3	110	94	AQ003188	AQ003188 RPII-11-4
22	85.8	0.3	105	105	AQ282107	AQ282107 RPII-11-94
23	85.4	0.3	103	94	AQ028649	AQ028649 CIT-HSP-2
24	85.6	0.3	110	33	AA442529	AA442529 zv68b02.f
25	85	0.3	101	39	AA835205	AA835205 ak64b01.s
26	85.2	0.3	106	38	AA812141	AA812141 OB48b02.s
27	85	0.3	110	64	AA083640	AA083640 xc49f02.x
28	84.6	0.3	100	35	AA564832	AA564832 nj22a06.s
29	84.8	0.3	104	105	AQ321855	AQ321855 RPII-11-11
30	84.4	0.3	102	36	AA654562	AA654562 nt75f10.s
31	84.4	0.3	102	84	B48088	B48088 RPII-11-4N6
32	84.6	0.3	103	108	AQ584425	AQ584425 RPII-11-4
33	84.6	0.3	108	84	B65160	B65160 CIT-HSP-201
34	84.2	0.3	107	24	H67040	H67040 yu68f01.r1
35	84.4	0.3	110	79	AA250394	AA250394 2822460.3
36	84.4	0.3	110	79	AA250394	AA250394 2822460.3
37	83.6	0.3	100	35	AA564832	AA564832 nj22a06.s
38	83.8	0.3	103	108	AQ534922	AQ534922 RPII-11-3
39	83.8	0.3	105	109	AQ637292	AQ637292 RPII-11-4
40	83.2	0.3	98	24	H67349	H67349 yu68f10.s1
41	83.4	0.3	103	107	AQ485214	AQ485214 RPII-11-2
42	83.2	0.3	104	108	AQ544583	AQ544583 CITBI-EI-
43	83.2	0.3	107	33	AA385808	AA385808 EST99495
44	83.2	0.3	107	103	AQ240182	AQ240182 CIT-HSP-2
45	83.4	0.3	109	30	AA243009	AA243009 zr25h02.s

ALIGNMENTS

```

RESULT 1
AA243009  AA243009  109 bp  mRNA  11-MAR-1998
LOCUS      zr25h02.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
DEFINITION cDNA clone IMAGE:664467 3', similar to contains Alu repetitive
            element; contains element LTR1 repetitive element ;, mRNA sequence.
ACCESSION  AA243009
VERSION     AA243009.1  GI:1873869
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 109)
AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wyllie,T., Waterston,R. and Wilson,R.
            WashU-NCI human EST Project
JOURNAL    Unpublished (1997)
COMMENT    On Dec 3, 1996 this sequence version replaced gi:1126869.

```

```

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1127 Std Error: 0.00
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 102.

FEATURES
    source
        1..109
            /organism="Homo sapiens"
            /db_xref="GDB:5426481"
            /db_xref="taxon:9606"
            /clone="IMAGE:664467"
            /clone_lib="Stratagene NT2 neuronal precursor 937230"
            /tissue_type="neuroepithelial cells"
            /dev_stage="Ntera-2 neuroepithelial cells"
            /lab_host="SOLR (kanamycin resistant)"
            /note="Organ: brain; Vector: pBluescript SK-; Site:1:
            EcoRI; Site2: XhoI; Cloned unidirectionally. Primer:
            Oligo dt. Uninduced, exponentially growing neuroepithelial
            cells (Ntera-2/ci.D1). Average insert size: 1.0 kb;
            Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGG
            3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT  19 a 30 c 30 g 30 t
ORIGIN
Query Match      0.3%; Score 93; DB 30; Length 109;
Best Local Similarity 90.8%; Pred. No. 0.22;
Matches 99; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 10849 GTATTTTATTAGATAGAGGGGTTTCCATGTTAGCCAGGATGCTCGATCTCTGACC 10908
            |||||
Db 1 GTATTTTATTAGATAGAGCGGGTTTCCACGCTGTAGCAGGATGCTCGATCTCTGACC 60
            |||||

Qy 10909 TCGTGATCCACCGCTTGGCTCCCAAGTCTGGGATTACAGCGTG 10957
            |||||
Db 61 TCGTGATCCGCGCTCCCAAGTCTGGGATTACAGCGTG 109
            |||||

RESULT 2
B17434/c  B17434  109 bp  DNA  04-JUN-1998
LOCUS      345K2.TVB CIT978SKAI Homo sapiens genomic clone A-345K02, genomic
DEFINITION survey sequence.
ACCESSION  B17434
VERSION     B17434.1  GI:2125183
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 109)
AUTHORS   Adams,M.D., Kelley,J.M., Rounsley,S.R. and Venter,J.C.
            Use of a BAC End Sequence Database for Sequence-Ready Map Building
            Unpublished (1997)
JOURNAL     Other-GSSs: 345K02.TP 345K02.TP
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdamas@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.

```


FEATURES
source

Location/Qualifiers
1. .106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="A-345R02"
/clone_lib="CIR978SK1"
/sex="Female"
/cell_type="Fibroblast"
/note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
Caltech Human BAC Library A1"

BASE COUNT
ORIGIN

24 a 30 c 31 g 24 t

Query Match 0.3%; Score 93; DB 84; Length 109;
Best Local Similarity 90.8%; Pred. No. 0.22; Mismatches 0; Gaps 0;
Matches 99; Conservative 0; Indels 10; Indels 0; Gaps 0;

QY 21813 TAGTAGACGGGGTTTACCATGTTGGTCAGGCTGGTCTGGAACCTCCTGACCTCAGGTG 21872

Db 109 TAGTTGAGACGGGGTTTACCATGTTGGTCAGGCTGGTCTCGAAGCTCCCGACCTCAGGTG 50

QY 21873 ATCTGCCACCTCAGGCTCCCAAGTGTGGGATTCACGATGAGCCCA 21921

Db 49 ATCCGCCACATCAGGCTCCCAAGTGTGGGATTCACGATGAGCCCA 1

RESULT 3

A1832832 105 bp mRNA EST 13-JUL-1999
LOCUS at72g09.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone
DEFINITION IMAGE:2377600 3' similar to contains Alu repetitive
element; contains element MER22 repetitive element ;, mRNA sequence.
ACCESSION A1832832 GI:5454812
VERSION EST.
KEYWORDS
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 105)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubouque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Dec 20, 1995 this sequence version replaced gi:1133644.
Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco.

FEATURES
source

Location/Qualifiers
1. .105
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2377600"
/clone_lib="Barstead colon HPLRB7"
/sex="male"
/dev_stage="adult, age 25"
/lab_host="DH10B (phage resistant)"
/note="Organ: colon; Vector: p7T3D-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACGATCTGAAGTGGAGCGGCCCTTTTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[5' AATTCACCTAGTAAT 3' and 5' ATTACTAGTG 3'], digested
with Not I and cloned into the Not I and Eco RI sites of

the modified p7T3 vector. Library constructed by Bob
Barstead."

BASE COUNT 17 a 35 c 27 g 26 t
ORIGIN

Query Match 0.3%; Score 92.2; DB 61; Length 105;

Best Local Similarity 92.4%; Pred. No. 0.28; Mismatches 0; Gaps 0;
Matches 97; Conservative 0; Indels 8; Indels 0; Gaps 0;

QY 5364 GAGATGGAGTTTCGCTCTTTGTTGCCAGGCTGGAGTGCATGCGGGATCTCGGCTCACC 5423

Db 1 GAGACAGAGTTTCGCTCTTTGTTGCCAGGCTGGAGTGCATGCGGGATCTCGGCTCACC 60

QY 5424 GCAACCTCTACCTCCAGCTTCAAGCAATTCCTCGCTCAGCCT 5468

Db 61 GCAACCTCCCTCCGGGTTCAAGCGATTCCTCGCTCAGCCT 105

RESULT 4

A1832832 106 bp mRNA EST 24-DEC-1997
LOCUS ag81a10.r1 Stragene hnt neuron (#937233) Homo sapiens cDNA clone
DEFINITION IMAGE:1140858 5' similar to contains Alu repetitive element; , mRNA
sequence.
ACCESSION A1832832 GI:2713610
VERSION EST.
KEYWORDS
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 106)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubouque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1397630.
Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 53.

FEATURES
source

Location/Qualifiers
1. .106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1140858"
/clone_lib="Stragene hnt neuron (#937233)"
/dev_stage="hnt neurons"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT.
Differentiated, post mitotic hnt neurons. Average insert
size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'
GAATTCGCGCAGGAG 3' -3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTTTTTT 3' "

BASE COUNT 19 a 29 c 29 g 29 t
ORIGIN

Query Match 0.3%; Score 91.6; DB 37; Length 106;

Best Local Similarity 91.5%; Pred. No. 0.32; Mismatches 0; Gaps 0;
Matches 97; Conservative 0; Indels 9; Indels 0; Gaps 0;

QY 10852 TTTTATTAGATGGGGTTTACCATTGTTAGCCAGATGCTCGATCTCCTGACCTCG 10911

```
Db 1 TTTTGTAGACAGAGGTTTACCGTGTACCCAGGATGCTCGATCTCTGACCTCG 60
Qy 10912 TGATCCACCGCTTGGCTCCCAAAGTCTGGGATTACAGCGTG 10957
Db 61 TGATCTGCCGCTCAGCTCCCAAAGTCTGGGATTACAGCGTG 106

RESULT 5
LOCUS AQ264176 106 bp DNA GSS 27-OCT-1998
DEFINITION CITBI-EI-2509A2.TF CITBI-EI Homo sapiens genomic clone 2509A2,
genomic survey sequence.
ACCESSION AQ264176
VERSION AQ264176.1 GI:3792743
KEYWORDS GSS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 106)
AUTHORS Adams M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K.,
Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and
Venter, J.C.
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT Other GSSs: CITBI-EI-2509A2.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2509A2"
/clone_lib="CITBI-EI"
/sex="male"
/cell_type="sperm"
/note="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"
BASE COUNT 25 a 30 c 34 g 17 t
ORIGIN

Query Match 0.3%; Score 91.6; DB 105; Length 106;
Best Local Similarity 91.5%; Pred. No. 0.32;
Matches 97; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 10867 GGGTTTACCATTGTTAGCCAGATGCTCGATCCCTGATCCACCGCTTT 10926
Db 106 GGGTTTACCATTGTTAGCCAGATGCTCGATCCCTGATCCACCGCTTC 47

Qy 10927 GCGCTCCCAAGTCTGGGATTACAGCGGTGAGCCACCGTGCCCGG 10972
Db 46 GGTCTCCCAAGTCTGGGATTACAGCGGTGAGATCTGCGCCCGG 1

RESULT 6
LOCUS AQ028426 109 bp DNA GSS 30-JUN-1998
DEFINITION CIT-HSP-2313G15.TF CIT-HSP Homo sapiens genomic clone 2313G15,
genomic survey sequence.
ACCESSION AQ028426
```

```
VERSION AQ028426.1 GI:3268648
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 109)
AUTHORS Adams M.D., Rounsley, S.D., Zhao, S., Field, C.E., Bass, S., Linher, K.,
Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H.,
Simon, M. and Venter, J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..109
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2313G15"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT 19 a 36 c 25 g 29 t
ORIGIN

Query Match 0.3%; Score 91.4; DB 94; Length 109;
Best Local Similarity 89.9%; Pred. No. 0.33;
Matches 98; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 15733 TTTTGTGAGACAGAGTCTCACTCTATCACCCAGGCTGGAGTGGCACAATCTC 15792
Db 1 TTGTTTCTGAGACGGACTCTCACTCTGTCAACCAGGCTGGAGTGGCAGACTGTG 60

Qy 15793 AGTCACATGCAACCTGCACCTCCCTGGGTTCAAGGATTCTCCTACCTAA 15841
Db 61 AGTCACATGCAACCTGCACCTCCCTGGGTTCAAGGATTCTCCTGCTCA 109

RESULT 7
LOCUS AQ535244 103 bp DNA GSS 18-MAY-1999
DEFINITION RPCI-11-317H22.TV RPCI-11 Homo sapiens genomic clone
RPCI-11-317H22, genomic survey sequence.
ACCESSION AQ535244
VERSION AQ535244.1 GI:4846934
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and
Venter, J.C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
```

9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@ig.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genet cs (inforesgen.com). BAC end search page: http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.

FEATURES

source

Location/Qualifiers

1. .103
/organism="Homo sapiens"
/db_xref="GDB:7621533"
/db_xref="taxon:9606"
/clone="RPCI-11-317H22"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"

BASE COUNT

ORIGIN

31 a 27 c 27 g 18 t
Query Match 0.3%; Score 90.2; DB 108; Length 103;
Best Local Similarity 92.2%; Pred. No. 0.46; Mismatches 8; Indels 0; Gaps 0;
Matches 95; Conservative 0;

QY 7027 CCAGCATTGGGAGCGCAAAAGGGGATCATTTGAGTCAGGAGTTCGAGACCCAGCC 7086
|||||
Db 1 CCAGCATTGGGAGCGCAAGAGCGGAGATCATTTGAGTCAGGAGTTCGAGACCCAGCC 60
QY 7087 TGCCCAACATGTTGAACCTCCATCTCTACTACAAAATACAAAA 7129
|||||
Db 61 TGGCCCAACATGTTGAACCCCGCTCTGCTATAAATACAAAA 103

RESULT 8

AA565533/c

LOCUS

DEFINITION AA565533 107 bp mRNA EST 08-SEP-1997
n42b11.s1 NCI-CGAP_GC2 Homo sapiens cDNA clone IMAGE:1016157 3', similar to contains Alu repetitive element,, mRNA sequence.

ACCESSION AA565533

VERSION AA565533.1 GI:2337172

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 107)

TITLE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

COMMENT Unpublished (1997)

On Sep 12, 1996 this sequence version replaced gi:1393355.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Stratagene, Inc., David B. Kriman, Ph.D.

CDNA Library Arraying: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/dbp/image/image.html

Insert Length: 1661 Std Error: 0.00

Seq primer: -40m13 fwd. Et from Amersham

FEATURES

source

High quality sequence stop: 87.

Location/Qualifiers

1. .107
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1016157"
/clone_lib="NCI-CGAP_GC2"
/tissue_type="germ cell tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dr. Bulk germ cell tumor. 5' adaptor sequence: 5' GAAATCGCAGAG 3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.2 kb."

BASE COUNT 22 a 34 c 26 g 25 t

ORIGIN

Query Match 0.3%; Score 89.4; DB 35; Length 107;
Best Local Similarity 89.7%; Pred. No. 0.55; Mismatches 11; Indels 0; Gaps 0;
Matches 96; Conservative 0;

QY 6217 TGTGTGTGCTGTAGTCCAGCTACTCAGAGGCTGGGCGCAGAGAGTTCGTTGAACCT 6276
|||||

Db 107 TGTGTGTGCTGTAAATCCAGCTACTCAGAGGCTGAGCGCAGAGAAATCACTTGAACCT 48
|||||

QY 6277 GCGAGCGGAGATTGCGATGCGGAGATCGCACCCGACCTCCAG 6323
|||||

Db 47 GGGAGGACGAGCTGTCAGTGTGAGATTGAGCCACTGCACCTCCAG 1
|||||

RESULT 9

AA078003/c

LOCUS

DEFINITION AA078003 105 bp mRNA EST 24-SEP-1999

7H12D08 Chromosome 7 HeLa cDNA Library Homo sapiens cDNA clone

AA078003

ACCESSION AA078003.1 GI:1837477

VERSION EST.

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 105)

AUTHORS Touchman, J.W., Bouffard, G.G., Weintrub, L.A., Idol, J.R., Wang, L., Robbins, C.M., Nussbaum, J.C., Lovett, M., and Green, E.D.

TITLE 2006 expressed-sequence tags derived from human chromosome 7-enriched cDNA libraries

JOURNAL Genome Res. 7 (3), 281-292 (1997)

MEDLINE 97228905

COMMENT On Apr 14, 1993 this sequence version replaced gi:693433.

Contact: Eric D. Green

Genome Technology Branch

National Human Genome Research Institute/NIH

49 Convent Dr., MSC4431, Building 49, Room 2A08, Bethesda, MD 20892

Tel: 3014020201

Fax: 3014024735

Email: egreen@nhgri.nih.gov

Plate: 12 row: D column: 08

Seq primer: -21M13 (ABI).

Location/Qualifiers

1. .105

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="7H12D08"

/sex="female"

/cell_line="HeLa cell line: ATCC"

/lab_host="E. coli strain DH5 alpha"

/note="Vector: pAMP10; cDNA was generated from cytoplasmic RNA using a mixture of random DNA hexamers and oligo(dT). From this pool of cDNA, human chromosome 7-enriched cDNA was isolated by direct cDNA selection using chromosome 7


```

RESULT 12
A0535244/c
LOCUS
DEFINITION
  A0535244      103 bp      DNA      GSS      18-MAY-1999
  RPCI-11-317H22.TV RPCI-11 Homo sapiens genomic clone
  RPCI-11-317H22, genomic survey sequence.
ACCESSION
  A0535244
VERSION
  A0535244.1 GI:4846934
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 103)
AUTHORS
  Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
  Venter,J.C.
  Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
  Map Building
  Unpublished (1997)
JOURNAL
  Contact: Shaying Zhao, William Nierman, Mark Adams
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: hbe@tigr.org
  Clones are derived from the human BAC library RPCI-11. For BAC
  library availability, please contact Pieter de Jong
  (pieter@dejong.med.buffalo.edu). Clones may be purchased from
  BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
  Research Genet cs (info@resgen.com). BAC end search page:
  http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.
  Seq primer: T7
  Class: BAC ends.
FEATURES
  source
  1..103
  /organism="Homo sapiens"
  /db_xref="GDB:7621533"
  /db_xref="taxon:9606"
  /clone="RPCI-11-317H22"
  /clone_lib="RPCI-11"
  /sex="Male"
  /cell_type="Lymphocytes"
  /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
  RPCI11 Human Male BAC Library"
BASE COUNT      31 a      27 c      27 g      18 t
ORIGIN
Query Match      0.3%; Score 88.6; DB 108; Length 103;
Best Local Similarity 91.3%; Pred. No. 0.69;
Matches 94; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 21801 TTTTGTATTTTATAGACACGGGTTTCACCATGTTGTGTGAGCTGGTCTGGAACCTCC 21860
|||||
Db 103 TTTTGTATTTATAGCAGACGGGTTTCACCATGTTGTGCCAGCTGGTCTCGAATCC 44
|||||

QY 21861 TGACCTCAGGTGATCGCCACCTCAGCCCTCCAAAGTCTGG 21903
|||||
Db 43 TGACCTCAAGTGATCGCCGCTTGTGCCCTCCAAAGTGTCTGG 1
|||||

RESULT 13
A0265749/c
LOCUS
DEFINITION
  A0265749      109 bp      DNA      GSS      27-OCT-1998
  CITBI-EI-2510E2.TR CITBI-EI Homo sapiens genomic clone 2510E2,
  genomic survey sequence.
ACCESSION
  A0265749
VERSION
  A0265749.1 GI:3791503
KEYWORDS
  GSS.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
  1 (bases 1 to 109)
AUTHORS
  Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
  Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
  Venter,J.C.
  Use of a random human BAC End Sequence Database for Sequence-Ready
  Map Building
  Unpublished (1998)
JOURNAL
  Other_GSSs: CITBI-EI-2510E2.TF
  Contact: Mark Adams
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: mdadams@tigr.org
  Clones are available from Research Genetics (info@resgen.com). BAC
  end search page:
  http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.
  Seq primer: M13 Reverse
  Class: BAC ends.
FEATURES
  Location/Qualifiers
  1..109
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="2510E2"
  /clone_lib="CITBI-EI"
  /sex="male"
  /cell_type="sperm"
  /note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;
  CalTech Human BAC Library D"
BASE COUNT      24 a      29 c      26 g      30 t
ORIGIN
Query Match      0.3%; Score 88.2; DB 105; Length 109;
Best Local Similarity 88.1%; Pred. No. 0.74;
Matches 96; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 9841 GCCAACATGGTGAACCTCTCTACTAAATAATACAAAAATAGCGGGCATGCTGGT 9900
|||||
Db 109 GCCAGCATGGTGAACCTCTCTCTACTAGAAATACAAAAATAGTCGCGGTGGCG 50
|||||

QY 9901 CAGCGCTGTATCCCGACTACTTGGGAGCTGAGACAGAGAATCGCTT 9949
|||||
Db 49 CATGCGCTGTATCCCGACTACTTGGGAAGCTGAGGCAGGAGCATCACTT 1
|||||

RESULT 14
A0004934
LOCUS
DEFINITION
  A0004934      102 bp      DNA      GSS      27-JUN-1998
  CIT-HSP-2292A10.TF CIT-HSP Homo sapiens genomic clone 2292A10,
  genomic survey sequence.
ACCESSION
  A0004934
VERSION
  A0004934.1 GI:3082379
KEYWORDS
  GSS.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 102)
AUTHORS
  Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
  Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
  Simon,M. and Venter,J.C.
  Use of a random BAC End Sequence Database for Sequence-Ready Map
  Building (1998)
JOURNAL
  Contact: Mark Adams
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: mdadams@tigr.org

```

Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13-21;
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1..102
/organism="Homo sapiens"
/db_xref="GDB:7151269"
/db_xref="taxon:9606"
/clone="2292A10"
/clone_lib="CIR-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII"

BASE COUNT 17 a 39 c 26 g 20 t
ORIGIN

Query Match 0.3%; Score 87.6; DB 94; Length 102;
Best Local Similarity 91.2%; Pred. No. 0.89;
Matches 93; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 5386 GCCCAGGCTGGAGTCAATGGCGGATCTCGGCTCACCGCAACCTCTACCTCCAGGTTTC 5445
Db 1 GCCCAGGCTGGAGTCTATGGCGGATCTCGGCTCACCGCAACCTCCAGGTTTC 60
QY 5446 AAGCAATCTCTGCTCAGCTCCCGAGTAGCTGGGATTAC 5487
Db 61 AAGCGACTCTCTGCTTAGCTCCCGAGTAGCTGGGATTAC 102

RESULT 15
AA807640

LOCUS AA807640 103 bp mRNA EST 05-MAR-1998
DEFINITION nx08b05.s1 NCI-CGAP_GC3 Homo sapiens cDNA clone IMAGE:1255473 3' similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION AA807640
VERSION AA807640.1 GI:2877108
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2151346.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 774 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 87.

FEATURES
source

Location/Qualifiers
1..103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1255473"
/clone_lib="NCI-CGAP_GC3"
/tissue_type="pooled germ cell tumors"

/lab_host="DH10B"
/note="vector: p7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from 3 pooled germ cell tumors, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T73 vector. Library is not normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 19 a 27 c 30 g 27 t
ORIGIN

Query Match 0.3%; Score 87.6; DB 38; Length 103;
Best Local Similarity 91.2%; Pred. No. 0.88;
Matches 93; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 10857 ATTAGAGATGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCTGACCTCTGATC 10916
Db 2 AGTAGAGATGGGGTTTCACCGTGTAGCCAGGATGGTCTCGATCTCTGACCTCTGATC 61
QY 10917 CACCCGCTTTGGCCCTCCCAAGTCTGGGATTACAGCGGTGA 10958
Db 62 CGCTCACCTCGCCCTCCCAAGTCTGGGATTACAGGTGTGA 103

Search completed: June 18, 2000, 04:51:11
Job time: 317638 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2000, 01:37:37 ; Search time 372.1 Seconds
(without alignments)
10130.869 Million cell updates/sec

Title: US-08-852-495c-2_COPY_28000_57000
Perfect score: 29001
Sequence: 1 GATAATTATTTAATTAT.....ACCTTTGATGAGACATGA 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2.6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2.6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2.6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn2.6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn2.6/ptodata/1/ina/6_COMB.seq:*
6: /cgn2.6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2.6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	80	0.3	105	4	US-08-481-658B-65
C 2	80	0.3	105	4	US-08-477-504A-65
C 3	80	0.3	105	4	US-08-486-756A-65
C 4	80	0.3	105	4	US-08-485-862B-65
C 5	80	0.3	105	5	US-08-787-739-65
C 6	79.4	0.3	105	4	US-08-481-658B-65
C 7	79.4	0.3	105	4	US-08-477-504A-65
C 8	79.4	0.3	105	4	US-08-486-756A-65
C 9	79.4	0.3	105	5	US-08-485-862B-65
C 10	79.4	0.3	105	5	US-08-787-739-65
C 11	65.4	0.2	84	3	US-08-454-557C-91
C 12	65.4	0.2	84	4	US-08-340-426D-91
C 13	65.4	0.2	84	4	US-08-450-673C-91
C 14	65.4	0.2	84	6	PCT-US95-17111A-91
C 15	62.8	0.2	84	3	US-08-454-557C-91
C 16	62.8	0.2	84	4	US-08-340-426D-91
C 17	62.8	0.2	84	4	US-08-450-673C-91
C 18	62.8	0.2	84	6	PCT-US95-17111A-91
C 19	60.8	0.2	85	3	US-08-454-557C-92
C 20	60.8	0.2	85	4	US-08-340-426D-92
C 21	60.8	0.2	85	4	US-08-450-673C-92
C 22	60.8	0.2	85	6	PCT-US95-17111A-92
C 23	60.4	0.2	78	3	US-08-454-557C-70
C 24	60.4	0.2	78	4	US-08-340-426D-70
C 25	60.4	0.2	78	4	US-08-450-673C-70
C 26	60.4	0.2	78	6	PCT-US95-17111A-70
C 27	58.8	0.2	78	3	US-08-454-557C-70

28	58.8	0.2	78	4	US-08-340-426D-70	Sequence 70, Appl
29	58.8	0.2	78	4	US-08-450-673C-70	Sequence 70, Appl
30	58.8	0.2	78	6	PCT-US95-17111A-70	Sequence 70, Appl
C 31	58	0.2	76	3	US-08-454-557C-69	Sequence 69, Appl
C 32	58	0.2	76	4	US-08-340-426D-69	Sequence 69, Appl
C 33	58	0.2	76	4	US-08-450-673C-69	Sequence 69, Appl
C 34	58	0.2	76	6	PCT-US95-17111A-69	Sequence 69, Appl
C 35	55.2	0.2	60	3	US-08-454-557C-57	Sequence 57, Appl
C 36	55.2	0.2	60	4	US-08-340-426D-57	Sequence 57, Appl
C 37	55.2	0.2	60	4	US-08-450-673C-57	Sequence 57, Appl
C 38	55.2	0.2	60	6	PCT-US95-17111A-57	Sequence 57, Appl
C 39	54.6	0.2	94	5	US-08-750-064-6	Sequence 6, Appl
C 40	53.6	0.2	60	3	US-08-454-557C-60	Sequence 60, Appl
C 41	53.6	0.2	60	4	US-08-340-426D-60	Sequence 60, Appl
C 42	53.6	0.2	60	4	US-08-450-673C-60	Sequence 60, Appl
C 43	53.6	0.2	60	6	PCT-US95-17111A-60	Sequence 60, Appl
C 44	51.6	0.2	76	3	US-08-454-557C-69	Sequence 69, Appl
C 45	51.6	0.2	76	4	US-08-340-426D-69	Sequence 69, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-65/c
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.3%; Score 80; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. No. 8.6e-10;

Matches 89; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 6099 ATCCGAGCACTTTGGAGGTCGAGCGAGCTGATCACGAGGTCAGAGTTCAAGACCAGC 6158
|||||
Db 105 ATCCGAGCACTTTGGAGGTCGAGCGAGCTGATCACGAGGTCAGAGTTCAAGACCAGC 46
|||||

QY 6159 CTGACCAAAATGATGAACCCCTGTCTCTACTAAAAATACAACA 6202
|||||
Db 45 CTGGCAATATGTTGAACCCCTGTCTCTACTAAAGATGTAAAAA 2

RESULT 2

US-08-477-504A-65/c
; Sequence 65, Application US/08477504A
; Patent No. 5972353

GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,504A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727

INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-477-504A-65

Query Match 0.3%; Score 80; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. NO. 8.6e-10;
Matches 89; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 6099 ATCCGAGCACTTTGGAGGTCGAGCGAGCTGATCACGAGGTCAGAGTTCAAGACCAGC 6158
|||||
Db 105 ATCCGAGCACTTTGGAGGTCGAGCGAGCTGATCACGAGGTCAGAGTTCAAGACCAGC 46
|||||

QY 6159 CTGACCAAAATGATGAACCCCTGTCTCTACTAAAAATACAACA 6202
|||||
Db 45 CTGGCAATATGTTGAACCCCTGTCTCTACTAAAGATGTAAAAA 2

RESULT 3

US-08-486-756A-65/c
; Sequence 65, Application US/08486756A
; Patent No. 5981711
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,756A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727

INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-486-756A-65

Query Match 0.3%; Score 80; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. NO. 8.6e-10;
Matches 89; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 6099 ATCCGAGCACTTTGGAGGTCGAGCGAGCTGATCACGAGGTCAGAGTTCAAGACCAGC 6158
|||||
Db 105 ATCCGAGCACTTTGGAGGTCGAGCGAGCTGATCACGAGGTCAGAGTTCAAGACCAGC 46
|||||

QY 6159 CTGACCAAAATGATGAACCCCTGTCTCTACTAAAAATACAACA 6202
|||||
Db 45 CTGGCAATATGTTGAACCCCTGTCTCTACTAAAGATGTAAAAA 2

RESULT 4

US-08-485-862B-65/c
; Sequence 65, Application US/08485862B
; Patent No. 5989638

GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court

:/ CITY: Tiburon
:/ STATE: California
:/ COUNTRY: USA
:/ ZIP: 94920
:/ MEDIUM TYPE: Floppy disk
:/ COMPUTER: IBM PC compatible
:/ OPERATING SYSTEM: PC-DOS/MS-DOS
:/ SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
:/ CURRENT APPLICATION DATA:
:/ APPLICATION NUMBER: US/08/485,862B
:/ FILING DATE: 07-JUN-1995
:/ CLASSIFICATION: 435
:/ PRIOR APPLICATION DATA:
:/ APPLICATION NUMBER: US 08/477,504
:/ FILING DATE: 07-JUN-1995
:/ APPLICATION NUMBER: US 08/260,190
:/ FILING DATE: 15-JUN-1994
:/ ATTORNEY/AGENT INFORMATION:
:/ NAME: Lauder, Leona L.
:/ REGISTRATION NUMBER: 30,863
:/ REFERENCE/DOCKET NUMBER: D-0021.3D
:/ TELECOMMUNICATION INFORMATION:
:/ TELEPHONE: 415-435-2034
:/ TELEFAX: 415-435-0727
:/ INFORMATION FOR SEQ ID NO: 65:
:/ SEQUENCE CHARACTERISTICS:
:/ LENGTH: 105 base pairs
:/ TYPE: nucleic acid
:/ STRANDEDNESS: single
:/ TOPOLOGY: linear
:/ MOLECULE TYPE: DNA (genomic)
:/ HYPOTHETICAL: NO
:/ ANTI-SENSE: NO
:/ US-08-485-862B-65

Query Match 0.3%; Score 80; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. No. 8.6e-10;
Matches 89; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 6099 ATCCAGCAGCTTTGGGAGGTCGAGGCGGCTGATCATCAGGAGTTCAGAGCCAGC 6158
Db 105 ATCCAGCAGCTTTGGGAGGCGGAGGCTGATCATCAGGAGTTCAGAGCCAGC 46

QY 6159 CTCACCAAAATGATGAACCCCTCTCTACTATAAAATACAAACA 6202
Db 45 CTGGCCAATATGTTGAACCCCTCTCTACTATAAAGATGTAAAAA 2

RESULT 5
US-08-787-739-65/c
:/ Sequence 65, Application US/08787739
:/ Patent No. 6027887
:/ GENERAL INFORMATION:
:/ APPLICANT: Zavada, Jan
:/ APPLICANT: Pastorekova, Silvia
:/ APPLICANT: Pastorek, Jaromir
:/ TITLE OF INVENTION: MN Gene and Protein
:/ NUMBER OF SEQUENCES: 96
:/ CORRESPONDENCE ADDRESS:
:/ ADDRESSEE: Leona L. Lauder
:/ STREET: 369 Pine Street, Suite 610
:/ CITY: San Francisco
:/ STATE: California
:/ COUNTRY: USA
:/ ZIP: 94104
:/ COMPUTER READABLE FORM:
:/ MEDIUM TYPE: Floppy disk
:/ COMPUTER: IBM PC compatible
:/ OPERATING SYSTEM: PC-DOS/MS-DOS
:/ SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
:/ CURRENT APPLICATION DATA:

:/ APPLICATION NUMBER: US/08/787,739
:/ FILING DATE: 24-JAN-1997
:/ PRIOR APPLICATION DATA:
:/ APPLICATION NUMBER: US 08/485,049
:/ FILING DATE: 07-JUN-1995
:/ PRIOR APPLICATION DATA:
:/ APPLICATION NUMBER: US 08/486,756
:/ FILING DATE: 07-JUN-1995
:/ PRIOR APPLICATION DATA:
:/ APPLICATION NUMBER: US 08/477,504
:/ FILING DATE: 07-JUN-1995
:/ PRIOR APPLICATION DATA:
:/ APPLICATION NUMBER: US 08/481,658
:/ FILING DATE: 07-JUN-1995
:/ PRIOR APPLICATION DATA:
:/ APPLICATION NUMBER: US 08/485,862
:/ FILING DATE: 07-JUN-1995
:/ PRIOR APPLICATION DATA:
:/ APPLICATION NUMBER: US 08/485,863
:/ FILING DATE: 07-JUN-1995
:/ PRIOR APPLICATION DATA:
:/ APPLICATION NUMBER: US 08/487,077
:/ FILING DATE: 07-JUN-1995
:/ ATTORNEY/AGENT INFORMATION:
:/ NAME: Lauder, Leona L.
:/ REGISTRATION NUMBER: 30,863
:/ REFERENCE/DOCKET NUMBER: D-0021.4
:/ TELECOMMUNICATION INFORMATION:
:/ TELEPHONE: 415-981-2034
:/ TELEFAX: 415-981-0332
:/ INFORMATION FOR SEQ ID NO: 65:
:/ SEQUENCE CHARACTERISTICS:
:/ LENGTH: 105 base pairs
:/ TYPE: nucleic acid
:/ STRANDEDNESS: double
:/ TOPOLOGY: linear
:/ MOLECULE TYPE: DNA (genomic)
:/ HYPOTHETICAL: NO
:/ ANTI-SENSE: NO
:/ US-08-787-739-65

Query Match 0.3%; Score 80; DB 5; Length 105;
Best Local Similarity 85.6%; Pred. No. 8.6e-10;
Matches 89; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 6099 ATCCAGCAGCTTTGGGAGGTCGAGGCGGCTGATCATCAGGAGTTCAGAGCCAGC 6158
Db 105 ATCCAGCAGCTTTGGGAGGCGGAGGCTGATCATCAGGAGTTCAGAGCCAGC 46

QY 6159 CTCACCAAAATGATGAACCCCTCTCTACTATAAAATACAAACA 6202
Db 45 CTGGCCAATATGTTGAACCCCTCTCTACTATAAAGATGTAAAAA 2

RESULT 6
US-08-481-658B-65
:/ Sequence 65, Application US/08481658B
:/ Patent No. 5955075
:/ GENERAL INFORMATION:
:/ APPLICANT: Zavada, Jan
:/ APPLICANT: Pastorekova, Silvia
:/ APPLICANT: Pastorek, Jaromir
:/ TITLE OF INVENTION: MN Gene and Protein
:/ NUMBER OF SEQUENCES: 86
:/ CORRESPONDENCE ADDRESS:
:/ ADDRESSEE: Leona L. Lauder
:/ STREET: 6 Mariposa Court
:/ CITY: Tiburon
:/ STATE: California
:/ COUNTRY: USA
:/ ZIP: 94920
:/ COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA: US/08/481,658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.3%; Score 79.4; DB 4; Length 105;
Best Local Similarity 84.8%; Pred. No. 1.2e-09;
Matches 89; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 10843 TTTTGTATTTTATTAGATGGGTTTACCATGTTAGCCAGGATGGTCTCGATCTC 10902
||||| || |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TTTTGTATTTTATTAGATGGGTTTACCATGTTAGCCAGGATGGTCTCGATCTC 60

QY 10903 CTGACCTCTGTATCCACCGCTTTGGCCCTCCCAAAGTGTCTGGAT 10947
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 CTGACCTCTGTATCCACCGCTTTGGCCCTCCCAAAGTGTCTGGAT 105

RESULT 7
US-08-477-504A-65
Sequence 65, Application US/08477504A
Patent No. 5972353
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-477-504A-65

Query Match 0.3%; Score 79.4; DB 4; Length 105;
Best Local Similarity 84.8%; Pred. No. 1.2e-09;
Matches 89; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 10843 TTTTGTATTTTATTAGATGGGTTTACCATGTTAGCCAGGATGGTCTCGATCTC 10902
||||| || |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TTTTGTATTTTATTAGATGGGTTTACCATGTTAGCCAGGATGGTCTCGATCTC 60

QY 10903 CTGACCTCTGTATCCACCGCTTTGGCCCTCCCAAAGTGTCTGGAT 10947
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 CTGACCTCTGTATCCACCGCTTTGGCCCTCCCAAAGTGTCTGGAT 105

RESULT 8
US-08-486-756A-65
Sequence 65, Application US/08486756A
Patent No. 5981711
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,756A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

```
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-486-756A-65

Query Match      0.3%; Score 79.4; DB 4; Length 105;
Best Local Similarity 84.8%; Pred. No. 1.2e-09;
Matches 89; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 10843 TTTTGTATTTTATTAGATGGGTTTCACCATGTTAGCCAGGATGCTCGATCTC 10902
Db      1 TTTTGTATCTTTTAGTAGAGACAGGTTTCACCATATTTGCCAGGCTGCTCTCAAACTC 60

QY 10903 CTGACCTCGTGATCCACCGCTTTGGCCTCCCAAAGTGTCTGGGAT 10947
Db      61 CTGACCTTGTGATCCACGACGCTCGGCCTCCCAAAGTGTCTGGGAT 105

RESULT 10
US-08-787-739-65
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-787-739-65

; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-486-756A-65

Query Match      0.3%; Score 79.4; DB 4; Length 105;
Best Local Similarity 84.8%; Pred. No. 1.2e-09;
Matches 89; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 10843 TTTTGTATTTTATTAGATGGGTTTCACCATGTTAGCCAGGATGCTCGATCTC 10902
Db      1 TTTTGTATCTTTTAGTAGAGACAGGTTTCACCATATTTGCCAGGCTGCTCTCAAACTC 60

QY 10903 CTGACCTCGTGATCCACCGCTTTGGCCTCCCAAAGTGTCTGGGAT 10947
Db      61 CTGACCTTGTGATCCACGACGCTCGGCCTCCCAAAGTGTCTGGGAT 105

RESULT 9
US-08-485-862B-65
; Sequence 65, Application US/08485862B
; Patent No. 5989838
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,862B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-485-862B-65

Query Match      0.3%; Score 79.4; DB 4; Length 105;
Best Local Similarity 84.8%; Pred. No. 1.2e-09;
Matches 89; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
```



```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450.673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-450-673C-91

Query Match      0.2%; Score 65.4; DB 4; Length 84;
Best Local Similarity 86.7%; Pred. No. 2e-06;
Matches 72; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 10875 CCATGTTAGCCAGGATGCTCGATCTCCTGACCTCGTGATCCACCGCTTTGGCGCTCCC 10934
      ||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1 CCATGTTATCATCAGCGTGGTGTGCAACTCCTGACCTCGTGATCCCGCGCTCAGCGCTCCC 60

QY 10935 AAAGTCTGGGATTACAGCGGTG 10957
      ||||| ||||| ||||| ||||| |||||
Db 61 AAAGTCTGGGATTACAGCGGTG 83

RESULT 14
PCT-US95-17111A-91
; Sequence 91, Application PC/TUS951711A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: PCT/US95/1711A
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340.426
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
```

```
;
; TOPOLOGY: both
PCT-US95-17111A-91

Query Match      0.2%; Score 65.4; DB 6; Length 84;
Best Local Similarity 86.7%; Pred. No. 2e-06;
Matches 72; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 10875 CCATGTTAGCCAGGATGCTCGATCTCCTGACCTCGTGATCCACCGCTTTGGCGCTCCC 10934
      ||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1 CCATGTTATCATCAGCGTGGTGTGCAACTCCTGACCTCGTGATCCCGCGCTCAGCGCTCCC 60

QY 10935 AAAGTCTGGGATTACAGCGGTG 10957
      ||||| ||||| ||||| ||||| |||||
Db 61 AAAGTCTGGGATTACAGCGGTG 83

RESULT 15
US-08-454-557C-91/C
; Sequence 91, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-454-557C-91

Query Match      0.2%; Score 62.8; DB 3; Length 84;
Best Local Similarity 85.4%; Pred. No. 7.9e-06;
Matches 70; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 6089 CATGACTGTAATCCAGCACACTTTGGAGGTCGAGGAGGCTATCATCAGAGGTCAGAGTT 6148
      || | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 83 CACGGTTGTAATCCAGCACACTTTGGAGGCTGAGGCGGCGGATCAGAGGTCAGAGTT 24

QY 6149 CAAGACCAGCCTGACCAAAATG 6170
      | | ||||| |||| | |||
Db 23 CGACACCAGCCTGATGAACATG 2
```

Search completed: June 18, 2000, 09:57:10
Job time: 334960 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2000, 09:50:49 ; Search time 17970.9 Seconds
(without alignments)
-1569.870 Million cell updates/sec

Title: US-08-852-495C-2_COPY_56000_85000
Perfect score: 29001
Sequence: 1 TCCCTTCAGTCTCCTCAAGGA.....TGGGGACCAAGTTTTTAAG 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba1.*
- 2: gb_ba2.*
- 3: gb_om.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_pil.*
- 8: gb_pl2.*
- 9: gb_pri.*
- 10: gb_pr2.*
- 11: gb_pr3.*
- 12: gb_ro.*
- 13: gb_sts.*
- 14: gb_sy.*
- 15: gb_un.*
- 16: gb_vi.*
- 17: em_fun.*
- 18: em_hum1.*
- 19: em_in.*
- 20: em_on.*
- 21: em_ov.*
- 22: em_ov.*
- 23: em_ov.*
- 24: em_pat.*
- 25: em_ph.*
- 26: em_pl.*
- 27: em_ro.*
- 28: em_sts.*
- 29: em_sy.*
- 30: em_un.*
- 31: em_vi.*
- 32: gb_htg1.*
- 33: gb_htg2.*
- 34: gb_in1.*
- 35: gb_in2.*
- 36: em_ba1.*
- 37: em_ba2.*
- 38: em_hum3.*
- 39: em_hum4.*
- 40: gb_pr4.*
- 41: gb_htg3.*
- 42: gb_htg4.*
- 43: gb_htg5.*
- 44: gb_htg6.*

- 45: gb_htg7.*
- 46: em_htg1.*
- 47: em_htg2.*
- 48: em_htg3.*
- 49: em_hum5.*
- 50: gb_pl3.*
- 51: gb_pr5.*
- 52: gb_htg8.*
- 53: gb_htg9.*
- 54: gb_htg10.*
- 55: gb_htg11.*
- 56: gb_htg12.*
- 57: gb_htg13.*
- 58: gb_htg14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	92	0.3	108	10	HSLDLRN2	X05250 Human LDL-r
2	92	0.3	108	10	HSLDLRN2	X05250 Human LDL-r
3	88.6	0.3	107	9	HUMALCE162	M87924 Human carc
4	88	0.3	108	11	HSU67803	U67803 Human small
5	84	0.3	103	9	HUMALCE221	M87896 Human carc
6	83.6	0.3	103	9	HUMALCE221	M87896 Human carc
7	83.8	0.3	104	9	HUMALCE272	M87899 Human carc
8	83	0.3	108	10	HSLDLRD1	X05249 Human LDL-r
9	83	0.3	108	10	HSLDLRD2	X05251 Human LDL-r
10	81.4	0.3	108	10	HSLDLRD1	X05249 Human LDL-r
11	81.4	0.3	108	10	HSLDLRD2	X05251 Human LDL-r
12	78.2	0.3	91	13	HUMUT8164A	L30244 Human STS U
13	78.4	0.3	108	10	HSLDLI12	X05248 Human LDL-r
14	77.8	0.3	108	11	HSU67803	U67803 Human small
15	77.8	0.3	110	9	HUMALCE43	M87900 Human carc
16	76	0.3	104	9	HUMALCE272	M87899 Human carc
17	76.2	0.3	108	11	HSU67808	U67808 Human small
18	76.4	0.3	108	11	HSU67808	U67808 Human small
19	76.4	0.3	110	9	HUMALCE43	M87900 Human carc
20	75.8	0.3	107	9	HUMALCE162	M87924 Human carc
21	75.4	0.3	106	13	G32743	G32743 A009P31 Hum
22	75.2	0.3	108	11	HSU67804	U67804 Human small
23	74.4	0.3	103	13	HS8IC8R	X57789 Human sequ
24	73.4	0.3	103	13	HS8IC8R	X57789 Human sequ
25	72	0.2	90	9	HUMDLRFL	R03555 Human low d
26	72	0.2	108	10	HSLDLI12	X05248 Human LDL-r
27	71.4	0.2	107	11	HSU67806	U67806 Human small
28	71.4	0.2	108	11	HSU67804	U67804 Human small
29	70.8	0.2	91	13	HUMUT8164A	L30244 Human STS U
30	70	0.2	107	11	HSU67806	U67806 Human small
31	70	0.2	108	9	HUMDL03M5	D16965 Human HepG2
32	70.2	0.2	110	11	HSU67807	U67807 Human small
33	70.2	0.2	110	11	HSU67807	U67807 Human small
34	69.4	0.2	97	9	HUMDLR2	M14180 Human low d
35	69.4	0.2	100	9	HUMALNSA	D45223 Human GALNS
36	68.8	0.2	80	9	HUMBRFAE	M36135 Human alpha
37	68.4	0.2	95	13	HUMUT8002B	L30176 Human STS U
38	68.4	0.2	102	13	G32906	G32906 A009W09 Hum
39	67.6	0.2	101	10	S79560	S79560 HRX (intron
40	67.8	0.2	108	9	HUMDL03M5	D16965 Human HepG2
41	67.2	0.2	84	5	AR051521	AR051521 Sequence
42	67.2	0.2	90	9	HUMDLRFL	R03555 Human low d
43	67.4	0.2	102	13	G32906	G32906 A009W09 Hum
44	67	0.2	100	13	HUMUT931A	L31299 Human STS U
45	66	0.2	100	10	HSLAS27	X91545 H.sapiens D

ALIGNMENTS

RESULT	1	
HSLDLRN2	108 bp	DNA
LOCUS	Human LDL-receptor gene	intron 14 fragment (normal gene).
DEFINITION	X05250	
ACCESSION	X05250.1	GI:34337
VERSION	Alu repetitive sequence;	low density lipoprotein receptor.
KEYWORDS	human.	
SOURCE	Human sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1	(bases 1 to 108)
AUTHORS	Horschmcke,B., Beisiegel,U., Dunning,A., Havinga,J.R., Williamson,R. and Humphries,S.	
TITLE	Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia	
JOURNAL	Eur. J. Biochem.	164 (1), 77-81 (1987)
MEDLINE	87161901	
COMMENT	See X05252 for deletion junction	
FEATURES	data kindly reviewed (07-DEC-1987) by HUMPHRIES S.	
source	Location/Qualifiers	
	1. .108	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
intron	1. .108	
	/note="intron XIV fragment"	
BASE COUNT	28 a	23 c 39 g 18 t
ORIGIN		
Query Match	0.3%	Score 92; DB 10; Length 108;
Best Local Similarity	90.7%	Pred. No. 2.2e-06;
Matches	98; Conservative	0; Mismatches 10; Indels 0; Gaps 0;
QY	10935	ACAAAGCTTAGCTGGGGCGTGGTGCACATGCCCTGTACTCCAGCTACTGGGGAGGCTGAG 10994
Db	1	ACAAAAATTACCGAGGCGTGGTGCAGGTGCTGTAACTCCAGCTACTCGGGAGGCTGAG 60
QY	10995	CGAGGAGAATTGCTTGAACTCGGAGGCGGAGGTTGCAGTGCAGCCGAG 11042
Db	61	CGAGGAGAATTGCTTGAACCCAGGAGGCGAGGTTGCAGTGCAGCCGAG 108
RESULT	2	
HSLDLRN2/c	108 bp	DNA
LOCUS	Human LDL-receptor gene	intron 14 fragment (normal gene).
DEFINITION	X05250	
ACCESSION	X05250.1	GI:34337
VERSION	Alu repetitive sequence;	low density lipoprotein receptor.
KEYWORDS	human.	
SOURCE	Human sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1	(bases 1 to 108)
AUTHORS	Horschmcke,B., Beisiegel,U., Dunning,A., Havinga,J.R., Williamson,R. and Humphries,S.	
TITLE	Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia	
JOURNAL	Eur. J. Biochem.	164 (1), 77-81 (1987)
MEDLINE	87161901	
COMMENT	See X05252 for deletion junction	
FEATURES	data kindly reviewed (07-DEC-1987) by HUMPHRIES S.	
source	Location/Qualifiers	
	1. .108	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
intron	1. .108	
	/note="intron XIV fragment"	
BASE COUNT	28 a	23 c 39 g 18 t

```

MEDLINE 97415756
REFERENCE 2 (bases 1 to 108)
AUTHORS Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The Children's Hospital of Philadelphia, 1004F Abramson Research Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES
    source 1..108
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="tscAlu2"
    repeat_region 1..108
        /note="sAlu"
        /rpt_family="Alu"
        /rpt_type="dispersed"
BASE COUNT 23 a 39 c 30 g 16 t
ORIGIN

Query Match 0.3%; Score 88; DB 11; Length 108;
Best Local Similarity 94.8%; Pred. No. 1.1e-05;
Matches 91; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 11565 GTAGATACGGGGTTTCACTTTGTTAAACAGGATGGTCTCGATCTCCCTGACCTCGTGATCG 11624
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 97 GTAGAGACGGGGTTTCACTTTGTTAGCAGGATGGTCTCGATCTCCCTGACCTCGTGATCC 38

QY 11625 GCCGCGCTCAGCCTCCCAAAGTGTGGGATTACAGG 11660
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 37 GCCGCGCTCAGCCTCCCAAAGTGTGGGATTACAGG 2

RESULT 5
HUMALCE221
LOCUS HUMALCE221 103 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
    Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Sinnott,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
    source 1..103
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /cell_line="Ntera2D1"
        /dev_stage="embryo"
        /sex="male"
        /tissue_type="carcinoma"
BASE COUNT 25 a 27 c 33 g 18 t
ORIGIN

Query Match 0.3%; Score 83.6; DB 9; Length 103;
Best Local Similarity 90.8%; Pred. No. 6.1e-05;
Matches 89; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 24288 CTGGAGTCTGTGCGACGTTCTCGGCTCACTGCAACCTCCGCCCTCAGGGTTCAAGCGAT 24347
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 103 CTGGAGTSCAATGCGACGATCTCGGCTCACTGCAACCTCCGCCCTCAGGGTTCAAGCGAT 44

QY 24348 TCTCTGCTCCGCGCTCCGAGTAGCTGAGATTACAGG 24385
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 43 TCTCTGCTCTAGCTTCCCGTGTAGCTGGGATTACAGG 6

RESULT 7
HUMALCE272/c
LOCUS HUMALCE272 104 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE272.
ACCESSION M87899
VERSION M87899.1 GI:174875
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
    Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 104)
AUTHORS Sinnott,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
    source 1..104
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /cell_line="Ntera2D1"
        /dev_stage="embryo"
        /sex="male"
        /tissue_type="carcinoma"
BASE COUNT 22 a 26 c 37 g 19 t
ORIGIN

Query Match 0.3%; Score 83.8; DB 9; Length 104;
Best Local Similarity 88.3%; Pred. No. 5.7e-05;
Matches 91; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```

```

RESULT 6
HUMALCE221/c
LOCUS HUMALCE221 103 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
    Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Sinnott,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
    source 1..103
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /cell_line="Ntera2D1"
        /dev_stage="embryo"
        /sex="male"
        /tissue_type="carcinoma"
BASE COUNT 25 a 27 c 33 g 18 t
ORIGIN

Query Match 0.3%; Score 83.6; DB 9; Length 103;
Best Local Similarity 90.8%; Pred. No. 6.1e-05;
Matches 89; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 24288 CTGGAGTCTGTGCGACGTTCTCGGCTCACTGCAACCTCCGCCCTCAGGGTTCAAGCGAT 24347
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 103 CTGGAGTSCAATGCGACGATCTCGGCTCACTGCAACCTCCGCCCTCAGGGTTCAAGCGAT 44

QY 24348 TCTCTGCTCCGCGCTCCGAGTAGCTGAGATTACAGG 24385
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 43 TCTCTGCTCTAGCTTCCCGTGTAGCTGGGATTACAGG 6

RESULT 7
HUMALCE272/c
LOCUS HUMALCE272 104 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE272.
ACCESSION M87899
VERSION M87899.1 GI:174875
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
    Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 104)
AUTHORS Sinnott,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
    source 1..104
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /cell_line="Ntera2D1"
        /dev_stage="embryo"
        /sex="male"
        /tissue_type="carcinoma"
BASE COUNT 22 a 26 c 37 g 19 t
ORIGIN

Query Match 0.3%; Score 83.8; DB 9; Length 104;
Best Local Similarity 88.3%; Pred. No. 5.7e-05;
Matches 91; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2000, 10:13:39 ; Search time 593.49 Seconds
(without alignments)
12225.683 Million cell updates/sec

Title: US-08-852-495C-2_COPY_56000_85000

Perfect score: 29001

Sequence: 1 TCCCTTCAGGCTCCAGGA.....TGGGACCAAAAGTTTTTAAG 29001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 433070

Minimum DB seq length: 10

Maximum DB seq length: 110

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_36:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	71	0.2	100	T24892	Human gene signatu
C 2	70.4	0.2	100	T24892	Human gene signatu
C 3	66.8	0.2	108	X12095	Human biallelic po
C 4	65.2	0.2	108	T26828	Human gene signatu
C 5	64.8	0.2	108	X12095	Human biallelic po
C 6	64.2	0.2	91	T25854	Human gene signatu
C 7	64.4	0.2	108	T26828	Human gene signatu
C 8	63.6	0.2	103	T26213	Human gene signatu
C 9	62.8	0.2	103	T20927	Human gene signatu
C 10	62.4	0.2	103	T20927	Human gene signatu
C 11	61	0.2	91	T25854	Human gene signatu
C 12	61	0.2	103	T26213	Human gene signatu
C 13	59.6	0.2	100	X12087	Human biallelic po
C 14	59.6	0.2	100	X12085	Human biallelic po
C 15	59.6	0.2	100	X12086	Human biallelic po
C 16	59	0.2	93	T22572	Human gene signatu
C 17	59.2	0.2	108	T25009	Human gene signatu
C 18	58	0.2	110	T26288	Human gene signatu
C 19	57.4	0.2	110	T26288	Human gene signatu
C 20	57	0.2	100	X12087	Human biallelic po
C 21	57	0.2	100	X12085	Human biallelic po
C 22	57	0.2	100	X12086	Human biallelic po
C 23	56	0.2	93	T25688	Human gene signatu
C 24	56.2	0.2	108	T25009	Human gene signatu
C 25	55	0.2	109	T23895	Human gene signatu
C 26	54.4	0.2	109	T23895	Human gene signatu
C 27	53.8	0.2	97	T26728	Human gene signatu
C 28	53.6	0.2	99	T20931	Human gene signatu
C 29	53.2	0.2	75	T22841	Human gene signatu
C 30	53.2	0.2	93	T22572	Human gene signatu
C 31	52.6	0.2	70	N60231	Normal chromosome
C 32	52.2	0.2	97	T26728	Human gene signatu
C 33	51.4	0.2	53	Q33621	Microsatellite seq
C 34	51.6	0.2	75	T22841	Human gene signatu

C 35	51.4	0.2	93	1	T24259	Human gene signatu
C 36	50.8	0.2	87	1	T21566	Human gene signatu
C 37	51	0.2	93	1	T24259	Human gene signatu
C 38	51	0.2	99	1	T20931	Human gene signatu
C 39	50.4	0.2	85	1	T26182	Human gene signatu
C 40	50.4	0.2	106	1	Q95210	Simple tandem repe
C 41	50	0.2	87	1	T21566	Human gene signatu
C 42	49.6	0.2	81	1	T24093	Human gene signatu
C 43	49.6	0.2	84	1	T25848	Human gene signatu
C 44	49.6	0.2	93	1	T25888	Human gene signatu
C 45	49.6	0.2	100	1	T25604	Human gene signatu

ALIGNMENTS

RESULT 1

T24892/c
ID T24892 standard; cDNA to mRNA; 100 BP.

AC T24892;

DT 05-NOV-1996 (first entry)

DE Human gene signature HUMG506998.

KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;

KW human; cloning; mapping; non-biased library; diagnosis; detection;

KW cell typing; abnormal cell function; ss.

OS Homo sapiens.

PN W09514772-Al.

PD 01-JUN-1995.

PR 11-NOV-1994; J01916.

PR 12-NOV-1993; JP-355504.

PA (MATS/) MATSUBARA K.

PA (OKUB/) OKUBO K.

PI Matsubara K., Okubo K;

DR WPI; 95-206931/27.

PT Identifying gene signatures in 3'-directed human cDNA library - e.g.

PT for diagnosis of abnormal cell function, by preparing cDNA that

PT reflects relative abundance of corresp. mRNA in specific human

PT tissues

PS Claim 1; Page 1720; 2245pp; Japanese.

CC A single-stranded DNA (or its complementary strand or the corresp.

CC double-stranded DNA) which comprises one of the 7837 "GS" sequences

CC given in T19001-T26837 and which is able to hybridise to part of

CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)

CC sequences were obtained from 3'-directed cDNA libraries prepared

CC from various human tissues; synthesis of cDNA was initiated from the

CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-

CC untranslated sequence is unique to a particular mRNA species, almost

CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library

CC is constructed so as to reflect accurately the relative abundance of

CC different mRNAs in the particular tissue from which it was derived.

CC The appearance frequency of a given GS in a cDNA library can be

CC determined (esp. using primers and probes derived from the GS

CC sequences) as a means of diagnosing abnormal cell function or for

CC recognising different cell types.

SQ Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;

Query Match

Best Local Similarity 0.2%; Score 71; DB 1; Length 100;

Matches 80; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 17560 TTTTTCCTTCAATAGAGTCTCGCTCTGTCCACCCAGGCTGAGTCGAGTCGGCCATCT 17619

Db 99 TTGTTTTCCTTCAATAGAGTCTCGCTCTGTCCACCCAGGCTGAGTCGAGTCGGCCATCT 40

QY 17620 CAGCTCACTGCAAGTCGCTCTCGCTCTGTCCACCCAGGCTGAGTCGAGTCGGCCATCT 17657

Db 39 CAGCTTATTCGAATTTTCCTCTCCACCCAGGTCACGGAT 2

RESULT 2

T24892

ID T24892 standard; cDNA to mRNA; 100 BP.


```

RESULT 6
T25854
ID T25854 standard; cDNA to mRNA; 91 BP.
AC T25854;
DT 22-OCT-1996 (first entry)
DE Human gene signature HUMGS08084.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
KW Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PI (OKUB/) OKUBO K.
PT Matsubara K. Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 144; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 91 BP; 18 A; 22 C; 28 G; 18 T;

Query Match 0.2%; Score 64.2; DB 1; Length 91;
Best Local Similarity 80.9%; Pred. No. 0.13;
Matches 72; Conservative 0; Mismatches 17; Indels 0; Gaps

QY 4066 ATCACTTAACTCAGGAGGCGAGGTTGCAGTGCAGTGAGTACACACCATTCGACTCCAG 4126
||||||| | ||||||| ||| ||||||||| ||| ||||||||| ||| |||||||
DB 2 ATCACTGACCTAGAGGCGAGGTTCAAGTGAGTGCAGTGCAGTGGCAGTCTCCTCGCTCCAG 61
||||||| | ||||||| ||| ||||||||| ||| ||||||||| ||| |||||||

QY 4126 CCTGGTGACAGTGTGCAGACTCTGTCTCA 4154
||| ||||||| ||||| |||||||
DB 62 CCTGGTGACAGCGTGAGANNCTGTCTCA 90

RESULT 7
T26828
ID T26828 standard; cDNA to mRNA; 108 BP.
AC T26828;
DT 14-NOV-1996 (first entry)
DE Human gene signature HUMGS09078.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
KW Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PI (OKUB/) OKUBO K.
PT Matsubara K. Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 144; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 91 BP; 18 A; 22 C; 28 G; 18 T;

```

PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 2182; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 18 A; 33 C; 23 G; 28 T;

Query Match 0.2%; Score 64.4; DB 1; Length 108;
Best Local Similarity 80.4%; Pred. No. 0.12;
Matches 74; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 11604 GATCTCTGACCTCGTATGTCGCGCCGCTCAGCTCCCAAGTGTGGATTACAGAGT 11663
DB 1 GATCTCTGACCTCGTATGTCGCGCCGCTCAGCTCCCAAGTGTGGATTACAGAGT 60

QY 11664 GATCCACTGCGCGCGCCGCTTTTTTTTTT 11695
DB 61 GAGCCACCACGCGCGCTGTTTATTTCTTAT 92

RESULT 8

ID T26213 standard; cDNA to mRNA; 103 BP.
AC T26213;
DT 13-NOV-1996 (first entry)
DE Human gene signature HUMGS08452.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 2029; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

SQ Sequence 103 BP; 33 A; 21 C; 25 G; 23 T;

Query Match 0.2%; Score 63.6; DB 1; Length 103;
Best Local Similarity 76.5%; Pred. No. 0.16;
Matches 78; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 10100 GATCTCTTGAGCTAGAAAGTTTGGGCGCGCTGAGTGTATGTCACCTGCACTCCA 10159
DB 1 GATCCTTGAGTCCAGGAGTTGGTTACAGTGAGCTATGATGCGCACCTGCACTCCA 60

QY 10160 GCCTGGGCAACAATGCAAAATCCTCTCTCAAAAACAAAACA 10201
DB 61 GCCTGGGCAACAATGCAAAATCCTCTCTTAAGAAAAA 102

RESULT 9

ID T20927 standard; cDNA to mRNA; 103 BP.
AC T20927;
DT 24-JUL-1996 (first entry)
DE Human gene signature HUMGS02180.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 758-759; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 103 BP; 22 A; 27 C; 21 G; 31 T;

Query Match 0.2%; Score 62.8; DB 1; Length 103;
Best Local Similarity 76.0%; Pred. No. 0.21;
Matches 76; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 17655 GATCTCTGCTCAGCCTTCCAGTAGCTGGGATTACAGATTGACGACCATGCCGAG 17714
DB 1 GATCTCCACCTCCACCTCCCAAGTAGCTGGGTACAGAGTGTGTGCCACCATGTCCAG 60

QY 17715 CTAAATTTTGTATTTTATAGAGAGACGGAATTTTACCATG 17754
DB 61 CTGATTTTNGTATTTTNNAGTAGGGACAGTATTTTCTCCATG 100

RESULT 10

ID T20927/c
ID T20927 standard; cDNA to mRNA; 103 BP.


```
Query Match          0.2%; Score 61; DB 1; Length 103;
Best Local Similarity 75.2%; Pred. No. 0.39;
Matches 76; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 17557 TTCCTTTTTCCTTTTGAATAGATCGCGCTCTACCCAGGCTGGAGTGCAGTGGCGAA 17616
    || ||||| || | | || ||||| ||||| ||||| ||||| || ||
Db 102 TTTTTCCTTTTAAAGACATGTTCTACTCTGTGCGCCAGGCTGGAGTGCAGTGGCGCA 43

QY 17617 TCTCAGCTACATGCAACGTCGCCCTCCCTGGGTTCAGTGTAT 17657
    || ||||| ||||| || | ||||| ||||| ||||| ||||| |||||
Db 42 TCATAGCTCACTGTAACACCAAACTCCTGGACTCAAGTGTAT 2

RESULT 13
X12087/c
ID X12087 standard; DNA; 100 BP.
AC X12087;
DE Human biallelic polymorphic DNA fragment EST98276a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
PI WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match          0.2%; Score 59.6; DB 1; Length 100;
Best Local Similarity 74.0%; Pred. No. 0.63;
Matches 74; Conservative 1; Mismatches 25; Indels 0; Gaps 0;

QY 4856 TGTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCTGAGGCGGCAGATCATCTGAGG 4915
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 100 TGTGACTCACACCTATAATCCTGGCACTTTAGGAGGCTTAGGAAGGAGGATGTTTGA 41

QY 4916 TCAGAAGTTCACAGACCGCTGCCCAACATGCGGAACCC 4955
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 40 CCAGGAGCTCAAGACCATCTCTGGAAACATAGCAAGACTC 1

RESULT 14
X12085/c
ID X12085 standard; DNA; 100 BP.
AC X12086 standard; DNA; 100 BP.
DE X12086;
DE 30-MAR-1999 (first entry)
KW Human biallelic polymorphic DNA fragment EST98276b.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
PI WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
```

```
AC X12085;
DE 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment EST98276c.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
PI WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 218; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 22 A; 25 C; 22 G; 30 T;

Query Match          0.2%; Score 59.6; DB 1; Length 100;
Best Local Similarity 74.0%; Pred. No. 0.63;
Matches 74; Conservative 1; Mismatches 25; Indels 0; Gaps 0;

QY 4856 TGTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCTGAGGCGGCAGATCATCTGAGG 4915
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 100 TGTGACTCACACCTATAATCCTGGCACTTTGGAGGCTTAGGAAGGAGGATGTTTGA 41

QY 4916 TCAGAAGTTCACAGACCGCTGCCCAACATGCGGAACCC 4955
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 40 CCAGGAGCTCAAGACCATCTCTGGAAACATAGCAAGACTC 1

RESULT 15
X12086/c
ID X12086 standard; DNA; 100 BP.
DE X12086;
DE 30-MAR-1999 (first entry)
KW Human biallelic polymorphic DNA fragment EST98276b.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
PI WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
```

PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match 0.2%; Score 59.6; DB 1; Length 100;
Best Local Similarity 74.0%; Pred. No. 0.63;
Matches 74; Conservative 1; Mismatches 25; Indels 0; Gaps 0;
QY 4856 TGTGGTCTACACCTGTATCCCGAGCACTTTGGAGGCTGAGCGGCGCATCTGTGAGG 4915
DB 100 TGTGACTCACACCTATATCTTGGCACTTTAGGAGGCTAGGAAGGAGGATTGTTGAAA 41
QY 4916 TCAGAGTTCCAGACCAGCTTGCCCAACATGGCGAACC 4955
DB 40 CCAGGAGCTCAAGACCATCTCTGGGAACATAGCAAGACTC 1

Search completed: June 18, 2000, 18:17:42
Job time: 364773 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2000, 04:51:11 ; Search time 8511.85 Seconds
(without alignments)
13809.868 Million cell updates/sec

Title: US-08-852-495C-2_COPY_56000_85000
Perfect score: 29001
Sequence: 1 TCCCTTCAGGCTCTCCAAGGA.....TGGGGACCAAAAGTTTTTAAAG 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : EST:*

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result % Query

SUMMARIES

Query Match	0.3%	Score 88.8;	DB 42;	Length 100;
Best Local Similarity	93.0%;	pred. No. 0.2;		
Matches 93;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;
Oy	22006	AAGACCTGGTGAAGGCAACTCTGTGTGCACAAGGCCCGGTTCCTTGCGTCC	22065	

Db 100 AAGAGCCTGTGGAGCAGCGGACTACTGTGCAAGAACGACGCGGTCTCTGGCTCC 41

QY 22066 TTAAACTCAACAAGAGCAGCCTCGGGAAGCAAGC 22105

Db 40 TTAAACTCAACAAGAGCAGCGGCTCGGGGAACCAAGC 1

RESULT 5

LOCUS AA835205 101 bp mRNA EST 23-FEB-1998

DEFINITION ak44h01.sl Barslead pancreas HPLRB1 Homo sapiens cDNA clone IMAGE:1412689 3' similar to contains Alu repetitive element; contains element KER repetitive element ;, mRNA sequence.

ACCESSION AA835205

VERSION AA835205.1 GI:2908933

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R. WashU-NCI human EST Project

TITLE Unpublished (1997)

JOURNAL On Nov 29, 1993 this sequence version replaced gi:636191.

COMMENT Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40m13 fwd. ET from Amersham.

FEATURES

Location/Qualifiers

1..101

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1412689"

/clone_lib="Barstead pancreas HPLRB1"

/sex="female"

/dev_stage="adult, 34 years"

/lab_host="DH10B"

/note="Organ: pancreas; vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACGAATCTGAAGTGGGAGCGCGCGCTTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [AATTCGATCGCTTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed by Bob Barstead."

14 a 36 c 27 g 24 t

BASE COUNT

ORIGIN

Query Match 0.3%; Score 88.8; DB 39; Length 101;

Best Local Similarity 93.0%; Pred. No. 0.2;

Matches 93; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 11402 GAGCGGAGTCTTGCTGTGGCCAGCGTGGAGTGCAGTGTGTGATCTCGGCTCACTG 11461

Db 2 GAGCGGAGTCTCACTGTGTGGCCAGCGTGGAGTGCAGTGTGTGATCTCGGCTCACTG 61

QY 11462 CAAGCTCCGCTCCCGGATTCACGCCATTCCTCGCTCA 11501

Db 62 CAAGCTCCGCTCCCGGATTCACGCCATTCCTCGCTCA 101

RESULT 6

B48914/c

LOCUS B48914 103 bp DNA GSS 08-APR-1999

DEFINITION RPCI11-4A12.TP RPCI-11 Homo sapiens genomic clone RPCI-11-4A12, genomic survey sequence.

ACCESSION B48914

VERSION B48914.1 GI:2601151

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and Venter, J.C.

TITLE Use of BAC End Sequences for Sequence-Ready Map Building

JOURNAL Unpublished (1997)

COMMENT Contact: Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: mdamas@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html

Seq primer: SP6

Class: BAC ends.

FEATURES

Location/Qualifiers

1..103

/organism="Homo sapiens"

/db_xref="GDB:7501163"

/db_xref="taxon:9606"

/clone="RPCI-11-4A12"

/clone_lib="RPCI-11"

/sex="Male"

/cell_type="Lymphocytes"

/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"

30 a 28 c 30 g 15 t

BASE COUNT

ORIGIN

Query Match 0.3%; Score 88.6; DB 84; Length 103;

Best Local Similarity 91.3%; Pred. No. 0.21;

Matches 94; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 11555 TGTATTTTGTAGATAGCGGGTTTCACTTTGTTAACAGGATGCTCTGATCTCTGAC 11614

Db 103 TGTATTTTGTAGATAGCGGGTTTCACTTTGTTAACAGGATGCTCTGATCTCTGAC 44

QY 11615 CTCGTGATCGCGCCCTCAGCCTCCCAAGTGTCTGGATTAC 11657

Db 43 CTCGTGATCGCGCCCTCAGCCTCCCAAGTGTCTGGCTTAC 1

RESULT 7

LOCUS B17434 109 bp DNA GSS 04-JUN-1998

DEFINITION B17434.345K2.TVB CIT978SK1 Homo sapiens genomic clone A-345K02, genomic survey sequence.

ACCESSION B17434

VERSION B17434.1 GI:2125183

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 109)

AUTHORS Adams,M.D., Kelley,J.M., Rounsley,S.R. and Venter,J.C.
TITLE Use of a BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other-GSSs: 345K02.TP 345K02.TPB
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadam@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.

FEATURES source
1. .109
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="A-345K02"
/clone_lib="CIT978SK1"
/sex="Female"
/cell_type="Fibroblast"
/note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
CalTech Human BAC Library A1"
BASE COUNT 24 a 30 c 31 g 24 t
ORIGIN

Query Match 0.3%; Score 88.2; DB 84; Length 109;
Best Local Similarity 88.1%; Pred. No. 0.23;
Matches 96; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 7539 TGGCTCATGCTGTATCCAGCACCTTTGGGAGGCTGAGTGGATGATCACCTGAGGTT 7598
||||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TGGCTCATGCTGTATCCAGCACCTTTGGGAGGCTGAGTGGGCGGATCACCTGAGGTC 60
QY 7599 GGCAGTTTGAGACGACCTGGCCCAACATGTTAAACCCCATGCTACTA 7647
||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GGGAGTTCGAGACGACCTGGCCACCATGTTGAACCCCGTCTCAACTA 109

RESULT 8
A0028426
LOCUS A0028426 109 bp DNA GSS 30-JUN-1998
DEFINITION CIT-HSP-2313G15.TF CIT-HSP Homo sapiens genomic clone 2313G15,
genomic survey sequence.
ACCESSION A0028426
VERSION A0028426.1 GI:3268648
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 109)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadam@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21

FEATURES source
1. .109
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2313G15"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT 19 a 36 c 25 g 29 t
ORIGIN

Query Match 0.3%; Score 88.2; DB 94; Length 109;
Best Local Similarity 88.1%; Pred. No. 0.23;
Matches 96; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 17561 TTTTGTGAAATAGAGTCTGCTGTCCACCCAGGCTGGAGTGGCGCAATCTC 17620
||||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TTGTTTTCGAGACGACTCTCACTCTGTCCACCCAGGCTGGAGTGGCGACAGTCTG 60
QY 17621 AGCTCACTGCAAGCTCGGCTCTCTGGTTCAGTGTCTCTCCGCTCA 17669
||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 AGCTCACTGCAAGCTCCACCTCTCTGGTTCAGGATTCCTCTGCTCA 109

RESULT 9
A0003188
LOCUS A0003188 110 bp DNA GSS 14-APR-1999
DEFINITION RPC111-1D10.TPN RPCI-11 Homo sapiens genomic clone RPCI-11-1D10,
genomic survey sequence.
ACCESSION A0003188
VERSION A0003188.1 GI:3030392
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 110)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venter,J.C.
TITLE Use of BAC End Sequences for Sequence-Ready Map Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadam@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

FEATURES source
1. .110
/organism="Homo sapiens"
/db_xref="GDB:750081"
/db_xref="taxon:9606"
/clone="RPCI-11-1D10"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC Library"
BASE COUNT 22 a 27 c 26 g 35 t

ORIGIN

Query Match 0.3%; Score 88.2; DB 94; Length 110;
Best Local Similarity 88.1%; Pred. No. 0.23; Mismatches 0; Gaps 0;
Matches 96; Conservative 0;

Qy 11551 TTTTGTATTTTGTAGATAGCGGGTTTCACTTTGTTTAAACCAGGATGCTCGCATCTCC 11610
LOCUS RPCI11-13414.TV RPCI-11 Homo sapiens genomic clone RPCI-11-13414,
Db 2 TTTTGTATTTTGTAGATAGCGGGTTTCACTTTGTTTAAACCAGGATGCTCGCATCTCT 61
genomic survey sequence.

Qy 11611 TGACCTCGTGATCGCCCGCTCAGCCCTCCCAAGTGTGGGATTACAG 11659
LOCUS RPCI11-13414.TV RPCI-11 Homo sapiens genomic clone RPCI-11-13414,
Db 62 TGACCTCGTGATCGCCCGCTCAGCCCTCCCAAGTGTGGGATTACAG 110
genomic survey sequence.

RESULT 10

AQ386882/c 110 bp DNA GSS 21-MAY-1999
DEFINITION RPCI11-13414.TV RPCI-11 Homo sapiens genomic clone RPCI-11-13414,
genomic survey sequence.

ACCESSION AQ386882
VERSION AQ386882.1 GI:4357905
KEYWORDS GSS.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 110)
AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and
Venter, J.C.

TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready

JOURNAL

COMMENT Map Building
Unpublished (1997)
Other_GSSs: RPCI11-13414.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetigr.org

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..110
/organism="Homo sapiens"
/db_xref="GDB:7551267"
/db_xref="taxon:9606"
/clones="RPCI-11-13414"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"

/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT 26 a 26 c 38 g 20 t
ORIGIN

Query Match 0.3%; Score 87.6; DB 106; Length 110;
Best Local Similarity 87.3%; Pred. No. 0.27; Mismatches 14; Indels 0; Gaps 0;
Matches 96; Conservative 0;

Qy 19521 GGGTTTACCAGTCGGCCAGGCTGCTCGAACCTCCTGACCTCAGCGGATCTGCCCGCC 19580
LOCUS RPCI11-13414.TV RPCI-11 Homo sapiens genomic clone RPCI-11-13414,
Db 110 GGGTTTACCAGTCGGCCAGGCTGCTCGAACCTCCTGACCTCAGCGGATCTGCCCGCC 110
genomic survey sequence.

Qy 19581 TCAGCCTCCCAAAAGTGTGATTACAGCGCTGAGCCACCAAGCCTGGCC 19630
LOCUS RPCI11-13414.TV RPCI-11 Homo sapiens genomic clone RPCI-11-13414,
Db 50 TCAGCCTCCCAAAAGTGTGATTACAGCGCTGAGCCACCAAGCCTGGCC 1
genomic survey sequence.

RESULT 11

AQ264176/c 106 bp DNA GSS 27-OCT-1998
DEFINITION CITBI-EI-2509A2.TF CITBI-EI Homo sapiens genomic clone 2509A2,
genomic survey sequence.

ACCESSION AQ264176
VERSION AQ264176.1 GI:3792743
KEYWORDS GSS.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 106)
AUTHORS Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and
Venter, J.C.

TITLE Use of a random human BAC End Sequence Database for Sequence-Ready

JOURNAL

COMMENT Map Building
Unpublished (1998)
Other_GSSs: CITBI-EI-2509A2.TR
Contact: Mark Adams

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CITBI-EI"
/sex="male"
/cell_type="sperm"

/note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"
BASE COUNT 25 a 30 c 34 g 17 t
ORIGIN

Query Match

Best Local Similarity 88.7%; Score 86.8; DB 105; Length 106;

Matches 94; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 11574 GGGTTTCACTTTGTTTAAACAGGATGCTCGATCTCTGATCTCTGATCGCGCCGCTC 11633
LOCUS RPCI11-13414.TV RPCI-11 Homo sapiens genomic clone RPCI-11-13414,
Db 106 GGGTTTCACTTTGTTTAAACAGGATGCTCGATCTCTGATCTCTGATCGCGCCGCTC 47
genomic survey sequence.

Qy 11634 AGCCTCCCAAGTGTGGGATTACAGGATGAGCCACATGCCGCCGG 11679
LOCUS RPCI11-13414.TV RPCI-11 Homo sapiens genomic clone RPCI-11-13414,
Db 46 GGTCTCCCAAAAGTGTGGGATTACAGGATGAGCCACATGCCGCCGG 1
genomic survey sequence.

RESULT 12

AA244245 110 bp mRNA EST 20-AUG-1997
DEFINITION nc07a04.s1 NC1-CCAP_Prl Homo sapiens cDNA clone IMAGE:1007406
similar to contains Alu repetitive element; mRNA sequence.

ACCESSION AA244245
VERSION AA244245.1 GI:1875104
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 110)
AUTHORS Chissoe, S., Dietrich, N., DuBuque, T., Favell, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Marais, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R., and Marra, M.

JOURNAL Unpublished (1997)
COMMENT On Jan 24, 1995 this sequence version replaced gi:634306.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuauqui, M.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Prepared by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 90.

FEATURES
source
1. .110
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1007406"
/clone_lib="NCI-CGAP_Prl"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected, histologically normal prostate epithelial cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

BASE COUNT 17 a 26 c 28 g 38 t 1 others
ORIGIN

Query Match 0.3%; Score 86.6; DB 30; Length 110;
Best Local Similarity 86.4%; Pred. No. 0.35;
Matches 95; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 17560 TTTTGTGTAATGAGTCTCGCTCTGTACCCAGGCTGGAGTGCAGTGGCGCAATCT 17619
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 TTTTGTGTAATGAGTCTCTGTATCTGTGCGGAGCTGGAGTGCAGTGGCGAGATCT 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 17620 CAGCTCAGTCAACGTCGCGCTCTCGGTTCAAGTATCTCTCGCTCA 17669
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 TGCTCAGTCAACCTCTGCTCTCTGGTTCAAGATCTCTCTGCTCA 110
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
LOCUS H67040 107 bp mRNA EST 27-OCT-1995
DEFINITION yu68c01.r1 weizmann Olfactory Epithelium Homo sapiens cDNA clone IMAGE:238944 5' similar to contains Alu repetitive element;; mRNA sequence.

ACCESSION H67040
VERSION H67040.1 GI:1025780
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 107)
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chissoe, S., Dietrich, N., DuBuque, T., Favell, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Marais, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R., and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT On Nov 29, 1993 this sequence version replaced gi:429999.
Contact: Willson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
High quality sequence stops: 101
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: M13RPI
High quality sequence stop: 101.

FEATURES
source
1. .107
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="GDB:3864328"
/db_xref="taxon:9606"
/clone="IMAGE:238944"
/clone_lib="Weizmann Olfactory Epithelium"
/sex="Female"
/tissue_type="olfactory epithelium"
/dev_stage="35 year old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: nose; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Olfactory epithelium, normal. Average insert size: 0.8 kb; Uni-ZAP XR Vector. Library constructed by N. Walker, D. Lancet, Weizmann Institute of Science. -5' adaptor sequence: 5' GAATTCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3' -"

BASE COUNT 24 a 37 c 20 g 24 t 2 others
ORIGIN

Query Match 0.3%; Score 85.8; DB 24; Length 107;
Best Local Similarity 86.9%; Pred. No. 0.44;
Matches 93; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 10943 TAGCTGGCGTGTGGCACATGCTGTAGTCCAGCTACTGGGAGGCTGAGCAGGAGA 11002
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 107 TAGCTGGGTGTGTAGCACATGCTGTATTCCNAGCTACTCAGNAGGCTGAGGTAGAGA 48
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 11003 ATGCTTGAACCTCGGAGCGGAGGTTGCAGTGCAGGAGATTCGCG 11049
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47 ATCGCTGACCCAGGAGGAGGTTGCAGTGCAGTGCAGATTGTC 1

RESULT 14
LOCUS B65160 108 bp DNA GSS 21-JUN-1998
DEFINITION CIT-HSP-2017G2.TRB CIT-HSP Homo sapiens genomic clone 2017G2, genomic survey sequence.

ACCESSION B65160
VERSION B65160.1 GI:2639138
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 108)
AUTHORS Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K.,

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2000, 09:57:10 ; Search time 372.13 Seconds
(without alignments)
10130.052 Million cell updates/sec

Title: US-08-852-495C-2_COPY_56000_85000
Perfect score: 29001
Sequence: 1 TCCCTTCAGGTCCTCCAAGGA.....TGGGACCAAAAGTTTTTAAG 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83.2	0.3	105	4	US-08-481-658B-65
2	83.2	0.3	105	4	US-08-477-504A-65
3	83.2	0.3	105	4	US-08-486-756A-65
4	83.2	0.3	105	4	US-08-485-862B-65
5	83.2	0.3	105	5	US-08-787-739-65
6	81.6	0.3	105	4	US-08-481-658B-65
7	81.6	0.3	105	4	US-08-477-504A-65
8	81.6	0.3	105	4	US-08-486-756A-65
9	81.6	0.3	105	5	US-08-787-739-65
10	81.6	0.3	105	5	US-08-485-862B-65
11	67.2	0.2	84	3	US-08-454-557C-91
12	67.2	0.2	84	4	US-08-340-426D-91
13	67.2	0.2	84	4	US-08-450-673C-91
14	67.2	0.2	84	6	PCT-US95-17111A-91
15	62.2	0.2	84	3	US-08-454-557C-91
16	62.2	0.2	84	4	US-08-340-426D-91
17	62.2	0.2	84	4	US-08-450-673C-91
18	62.2	0.2	84	6	PCT-US95-17111A-91
19	58.8	0.2	78	3	US-08-454-557C-70
20	58.8	0.2	78	4	US-08-340-426D-70
21	58.8	0.2	78	4	US-08-450-673C-70
22	58.8	0.2	78	6	PCT-US95-17111A-70
23	54.6	0.2	85	3	US-08-454-557C-92
24	54.4	0.2	85	4	US-08-454-557C-92
25	54.6	0.2	85	4	US-08-340-426D-92
26	54.4	0.2	85	4	US-08-340-426D-92
27	54.6	0.2	85	4	US-08-450-673C-92

Sequence 92, Appl
Sequence 92, Appl
Sequence 92, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 66, Appl
Sequence 66, Appl
Sequence 66, Appl
Sequence 66, Appl
Sequence 66, Appl
Sequence 70, Appl
Sequence 70, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-65
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.3%; Score 83.2; DB 4; Length 105;
Best Local Similarity 87.5%; Pred. No. 1.4e-09;

Matches 91; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 27950 TTTTGTATTTTATTAGACAGGGTTTCACTATGTTGGCCAGGCTGATCTCAAACTCC 28009

Db 2 TTTTACATCTTTAGTAGACAGGGTTTCACTATGTTGGCCAGGCTGCTCTCAAACTCC 61

Qy 28010 TGACCTCATGATCCGCCCTGCGCTTGGCCCTCTCAAAAGTGTCTGGAT 28053

Db 62 TGACCTTGATCCACCAGCCTCGGCCCTCCCAAAAGTGTCTGGAT 105

RESULT 2

US-08-477-504A-65

; Sequence 65, Application US/08477504A

; Patent No. 5972353

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,504A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3D

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-477-504A-65

Query Match 0.3%; Score 83.2; DB 4; Length 105;

Best Local Similarity 87.5%; Pred. No. 1.4e-09;

Matches 91; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 27950 TTTTGTATTTTATTAGACAGGGTTTCACTATGTTGGCCAGGCTGATCTCAAACTCC 28009

Db 2 TTTTACATCTTTAGTAGACAGGGTTTCACTATGTTGGCCAGGCTGCTCTCAAACTCC 61

Qy 28010 TGACCTCATGATCCGCCCTGCGCTTGGCCCTCTCAAAAGTGTCTGGAT 28053

Db 62 TGACCTTGATCCACCAGCCTCGGCCCTCCCAAAAGTGTCTGGAT 105

RESULT 3

US-08-486-756A-65

; Sequence 65, Application US/08486756A

; Patent No. 5981711

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,756A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-486-756A-65

Query Match 0.3%; Score 83.2; DB 4; Length 105;

Best Local Similarity 87.5%; Pred. No. 1.4e-09;

Matches 91; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 27950 TTTTGTATTTTATTAGACAGGGTTTCACTATGTTGGCCAGGCTGATCTCAAACTCC 28009

Db 2 TTTTACATCTTTAGTAGACAGGGTTTCACTATGTTGGCCAGGCTGCTCTCAAACTCC 61

Qy 28010 TGACCTCATGATCCGCCCTGCGCTTGGCCCTCTCAAAAGTGTCTGGAT 28053

Db 62 TGACCTTGATCCACCAGCCTCGGCCCTCCCAAAAGTGTCTGGAT 105

RESULT 4

US-08-485-862B-65

; Sequence 65, Application US/08485862B

; Patent No. 5989838

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481.658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.3%; Score 81.6; DB 4; Length 105;
Best Local Similarity 86.5%; Pred. No. 3.1e-09;
Matches 90; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 24972 ATCCAGCTCTTTGGAGGCGCTAGCGGTGGATCACAGGTCAGGAGTTCAAGACGAGC 25031
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 105 ATCCAGCAGCTTTGGAGGCGCGAGCTGGTGGATCACAAAGGTCAGGAGTTTGAGAGCAGC 46

Qy 25032 CTCGCCAAGATGGTGAATCCGCTCTCTACTAAAAGTATAAAA 25075
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 45 CTGCCAATATGGTGAACCCCTGCTCTACTAAAGATGTAATAA 2

RESULT 7
US-08-477-504A-65/c
Sequence 65, Application US/08/477504A
Patent No. 5972553
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477.504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-477-504A-65

Query Match 0.3%; Score 81.6; DB 4; Length 105;
Best Local Similarity 86.5%; Pred. No. 3.1e-09;
Matches 90; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 24972 ATCCAGCTCTTTGGAGGCGCTAGCGGTGGATCACAGGTCAGGAGTTCAAGACGAGC 25031
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 105 ATCCAGCAGCTTTGGAGGCGCGAGCTGGTGGATCACAAAGGTCAGGAGTTTGAGAGCAGC 46

Qy 25032 CTCGCCAAGATGGTGAATCCGCTCTCTACTAAAAGTATAAAA 25075
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 45 CTGCCAATATGGTGAACCCCTGCTCTACTAAAGATGTAATAA 2

RESULT 8
US-08-486-756A-65/c
Sequence 65, Application US/08486756A
Patent No. 5981711
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486.756A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-486-756A-65

Query Match 0.3%; Score 81.6; DB 4; Length 105;
Best Local Similarity 86.5%; Pred. No. 3.1e-09;
Matches 90; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 24972 ATCCGAGCTCTTTGGAGGCGCTAGGCGGTGGATCAGAGGTCAGAGTTCAAGACCAGC 25031
DB 105 ATCCGAGCAGCTTTGGAGGCGCGAGGCTGGTGATCACAAGGTCAGAGTTTGAGAGCAGC 46
QY 25032 CTCGCCAAGATGGTGAATCCCGTCTCTACTATAAAGTATAAAAA 25075
DB 45 CTGGCCATATGTTGAACCCCTGCTCTACTATAAAGATGTAAAAA 2

RESULT 9
US-08-485-862B-65/c
Sequence 65, Application US/08485862B
Patent No. 5989838
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485.862B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-485-862B-65

Query Match 0.3%; Score 81.6; DB 4; Length 105;
Best Local Similarity 86.5%; Pred. No. 3.1e-09;
Matches 90; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 24972 ATCCGAGCTCTTTGGAGGCGCTAGGCGGTGGATCAGAGGTCAGAGTTCAAGACCAGC 25031
DB 105 ATCCGAGCAGCTTTGGAGGCGCGAGGCTGGTGATCACAAGGTCAGAGTTTGAGAGCAGC 46
QY 25032 CTCGCCAAGATGGTGAATCCCGTCTCTACTATAAAGTATAAAAA 25075
DB 45 CTGGCCATATGTTGAACCCCTGCTCTACTATAAAGATGTAAAAA 2

RESULT 10
US-08-787-739-65/c
Sequence 65, Application US/08787739
Patent No. 6027887
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 96
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 369 Pine Street, Suite 610
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/787,739
FILING DATE: 24-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/486,756
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/481,658
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,862
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,863
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,077
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.4
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-787-739-65

```
Query Match      0.3%; Score 81.6; DB 5; Length 105;
Best Local Similarity 86.5%; Pred. No. 3.1e-09;
Matches 90; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 24972 ATCCGAGCTCTTTGGGAGGCTAGCGGTGGATCAGAGTCTCAGGAGTTCAAGACGAGC 25031
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 105 ATCCGAGCACTTTGGGAGGCGGAGGCTGGTGATCACAAGTCTCAGGAGTTTGAGAGCAGC 46
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 25032 CTCGCCAAGATGGTGAATCCGCTCTTACTATAAAGTATAAAA 25075
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 45 CTGCCAATATGGTGAACCCCTGTCTCTACTAAAGATGTAATAA 2

RESULT 11
US-08-454-557C-91
; Sequence 91, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-454-557C-91

Query Match      0.2%; Score 67.2; DB 3; Length 84;
Best Local Similarity 90.0%; Pred. No. 4e-06;
Matches 72; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 11585 TGTTAACCAAGAGTGTCTCGATCTCTGACCTCGTGATCGCGCGCTCAGCCCTCCCAAA 11644
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4 TGTTCATCAGGCTGGTGTGCAACTCTCTGACCTCGTGATCGCGCGCTCAGCCCTCCCAAA 63
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 11645 GTGCTGGGATTACAGGAGTG 11664
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 64 GTGCTGGGATTACAAAGCGTG 83
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
US-08-340-426D-91
; Sequence 91, Application US/08340426D
; Patent No. 5948634
```

```
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-340-426D-91

Query Match      0.2%; Score 67.2; DB 4; Length 84;
Best Local Similarity 90.0%; Pred. No. 4e-06;
Matches 72; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 11585 TGTTAACCAAGAGTGTCTCGATCTCTGACCTCGTGATCGCGCGCTCAGCCCTCCCAAA 11644
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4 TGTTCATCAGGCTGGTGTGCAACTCTCTGACCTCGTGATCGCGCGCTCAGCCCTCCCAAA 63
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 11645 GTGCTGGGATTACAGGAGTG 11664
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 64 GTGCTGGGATTACAAAGCGTG 83
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-08-450-673C-91
; Sequence 91, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
```



```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-450-673C-91

Query Match 0.2%; Score 67.2; DB 4; Length 84;
Best Local Similarity 90.0%; Pred. No. 4e-06;
Matches 72; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 11585 TGTTAACGAGGTGCTCGATCTCCTGACCTCGTGATCGCGCGCTCAGCCTCCCAAA 11644
Db 4 TGTTCATCAGCGTGGTGTGCAACTCTGACCTCGTGATCGCGCGCTCAGCCTCCCAAA 63

QY 11645 GTCTGGGATTACAGGAGTG 11664
Db 64 GTCTGGGATTACAGCGTG 83

RESULT 14
PCT-US95-17111A-91
; Sequence 91, Application PC/TUS951711A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-454-557C-91

Query Match 0.2%; Score 62.2; DB 3; Length 84;
Best Local Similarity 84.3%; Pred. No. 5e-05;
Matches 70; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 24962 CAGCGCTATATATCCAGCTCTTTGGGAGGCGCTAGGCGGGTGATCAGAGTCAAGGAGTT 25021
Db 83 CAGCGTTGTATATCCAGCACTTTGGGAGGCTAGGCGGGTGATCAGAGTCAAGGAGTT 24

QY 25022 CAAGACCAGCCTCGCCCAAGATGG 25044
Db 23 CGACACCAGCCTGATGAACATGG 1
```

```
;
; TOPOLOGY: both
;
PCT-US95-17111A-91

Query Match 0.2%; Score 67.2; DB 6; Length 84;
Best Local Similarity 90.0%; Pred. No. 4e-06;
Matches 72; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 11585 TGTTAACGAGGTGCTCGATCTCCTGACCTCGTGATCGCGCGCTCAGCCTCCCAAA 11644
Db 4 TGTTCATCAGCGTGGTGTGCAACTCTGACCTCGTGATCGCGCGCTCAGCCTCCCAAA 63

QY 11645 GTCTGGGATTACAGGAGTG 11664
Db 64 GTCTGGGATTACAGCGTG 83

RESULT 15
US-08-454-557C-91/c
; Sequence 91, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-454-557C-91

Query Match 0.2%; Score 62.2; DB 3; Length 84;
Best Local Similarity 84.3%; Pred. No. 5e-05;
Matches 70; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 24962 CAGCGCTATATATCCAGCTCTTTGGGAGGCGCTAGGCGGGTGATCAGAGTCAAGGAGTT 25021
Db 83 CAGCGTTGTATATCCAGCACTTTGGGAGGCTAGGCGGGTGATCAGAGTCAAGGAGTT 24

QY 25022 CAAGACCAGCCTCGCCCAAGATGG 25044
Db 23 CGACACCAGCCTGATGAACATGG 1
```

Search completed: June 18, 2000, 17:46:45
Job time: 363135 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2000, 17:40:24 ; Search time 17971.1 seconds
(without alignments)
-1569.847 Million cell updates/sec

Title: US-08-852-495C-2_COPY_84000_113000
Perfect score: 29001
Sequence: 1 TCAAACTCCTGACCTCATGA.....ATAAATATCTTTAAATACC 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba1.*
- 2: gb_ba2.*
- 3: gb_om.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_p11.*
- 8: gb_p12.*
- 9: gb_p13.*
- 10: gb_p14.*
- 11: gb_p15.*
- 12: gb_ro.*
- 13: gb_sts.*
- 14: gb_sy.*
- 15: gb_un.*
- 16: gb_v1.*
- 17: em_fun.*
- 18: em_hum1.*
- 19: em_hum2.*
- 20: em_in.*
- 21: em_om.*
- 22: em_or.*
- 23: em_ov.*
- 24: em_pat.*
- 25: em_ph.*
- 26: em_pl.*
- 27: em_ro.*
- 28: em_sts.*
- 29: em_sy.*
- 30: em_un.*
- 31: em_v1.*
- 32: gb_htg1.*
- 33: gb_htg2.*
- 34: gb_in1.*
- 35: gb_in2.*
- 36: em_ba1.*
- 37: em_ba2.*
- 38: em_hum3.*
- 39: em_hum4.*
- 40: gb_pr4.*
- 41: gb_htg3.*
- 42: gb_htg4.*
- 43: gb_htg5.*
- 44: gb_htg6.*

- 45: gb_htg7.*
- 46: em_htg1.*
- 47: em_htg2.*
- 48: em_htg3.*
- 49: em_hum5.*
- 50: gb_p13.*
- 51: gb_pr5.*
- 52: gb_htg8.*
- 53: gb_htg9.*
- 54: gb_htg10.*
- 55: gb_htg11.*
- 56: gb_htg12.*
- 57: gb_htg13.*
- 58: gb_htg14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	96.6	0.3	107	9	HUMALCE162	M87924 Human carci
2	89	0.3	108	11	HSU67803	U67803 Human small
C 3	87.4	0.3	108	10	HSIDLNR2	X05250 Human LDL-r
4	86.8	0.3	108	10	HSIDLNR2	X05250 Human LDL-r
C 5	85.4	0.3	107	9	HUMALCE162	M87924 Human carci
6	81.4	0.3	103	9	HUMALCE221	M87896 Human carci
C 7	81.4	0.3	103	9	HUMALCE221	M87896 Human carci
C 8	81	0.3	108	11	HSU67803	U67803 Human small
9	80	0.3	108	10	HSIDLNR1	X05249 Human LDL-r
C 10	80	0.3	108	10	HSIDLNR2	X05251 Human LDL-r
C 11	79.4	0.3	108	10	HSIDLNR1	X05249 Human LDL-r
12	79.4	0.3	108	10	HSIDLNR2	X05251 Human LDL-r
13	78.2	0.3	108	10	HSIDLNR2	X05248 Human LDL-r
14	77.8	0.3	108	11	HSU67804	U67804 Human small
15	76	0.3	104	9	HUMALCE272	M87899 Human carci
16	75.4	0.3	97	9	HUMDLRA2	M14180 Human low d
17	74.8	0.3	110	9	HUMALCE43	M87900 Human carci
C 18	73.8	0.3	103	13	HS8IC8R	X57789 Human sequ
19	73	0.3	108	11	HSU67808	U67808 Human small
C 20	72.6	0.3	106	13	G32743	G32743 A009P31 Hum
C 21	72.8	0.3	110	11	HSU67807	U67807 Human small
C 22	72.4	0.2	91	13	HUMUT8164A	L30244 Human STS U
C 23	72.4	0.2	108	10	HSIDLNR1	X05248 Human LDL-r
C 24	72	0.2	90	9	HUMDLRFL	K03555 Human low d
C 25	72.2	0.2	108	9	HUMDLRFL	D16965 Human HepG2
C 26	71.4	0.2	107	11	HSU67806	U67806 Human small
C 27	71.4	0.2	108	11	HSU67804	U67804 Human small
C 28	71.6	0.2	110	11	HSU67807	U67807 Human small
C 29	71.2	0.2	103	13	HS8IC8R	X57789 Human sequ
C 30	70.4	0.2	80	9	HUMBRKFAE	M36135 Human alpha
C 31	69.8	0.2	108	11	HSU67808	U67808 Human small
C 32	68.4	0.2	95	13	HUMUT8002B	L30176 Human STS U
C 33	67.6	0.2	84	5	AR051521	AR051521 Sequence
C 34	67.2	0.2	80	9	HUMBRKFAE	M36135 Human alpha
C 35	67.4	0.2	94	9	HUMHGAL	M13479 Human alpha
C 36	67.4	0.2	95	13	HUMUT8002B	L30176 Human STS U
C 37	66.6	0.2	91	13	HUMUT8164A	L30244 Human STS U
C 38	66.8	0.2	94	9	HUMHGAL	M13479 Human alpha
C 39	66	0.2	97	9	HUMDLRAL	M14178 Human low d
C 40	66.2	0.2	108	9	HUMDLR1	D16965 Human HepG2
C 41	66.2	0.2	108	13	G43535	G43535 WIAT-2393-S
C 42	65.6	0.2	97	9	HUMDLR1	M14179 Human famil
C 43	65.2	0.2	79	10	S73203	S73203 ALL-1 (tand
C 44	64.8	0.2	76	11	AF032287	AF032287 Eulemur m
C 45	65	0.2	107	11	HSU67806	U67806 Human small

ALIGNMENTS

```
RESULT 1
LOCUS HUMALCE162 107 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE162.
ACCESSION M87924
VERSION M87924.1 GI:174871
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 107)
AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
    source
        1..107
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /cell_line="Ntera2D1"
            /dev_stage="embryo"
            /sex="male"
            /tissue_type="carcinoma"
BASE COUNT 28 a 30 c 35 g 14 t
ORIGIN

Query Match 0.3%; Score 96.6; DB 9; Length 107;
Best Local Similarity 96.1%; Pred. No. 2.3e-09;
Matches 99; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9477 GCGAGGAATGGCTGACCGGAGCGGAGCTTGCGAGTGCAGCGGAGATCGCCACG 9536
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 5 GCGAAGATGGCTGACCGGAGCGGAGCTTGCGAGTGCAGCGGAGATCGCCACT 64
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 9537 GCACCTCCAGCTGGTGACAGCGAGACTCGTCTCAAAAAA 9579
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 65 GCACCTCCAGCTGGTGACAGCGAGACTCGTCTCAAAAAA 107
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 2
LOCUS HSU67803 108 bp RNA PRI 01-AUG-1997
DEFINITION Human small cytoplasmic Alu transcript.
ACCESSION U67803
VERSION U67803.1 GI:2289917
KEYWORDS Alu.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE cDNAs derived from primary and small cytoplasmic Alu (scAlu)
transcripts
JOURNAL J. Mol. Biol. 271 (2), 222-234 (1997)
MEDLINE 97415756
REFERENCE 2 (bases 1 to 108)
AUTHORS Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
Children's Hospital of Philadelphia, 1004F Abramson Research
Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES
    source
        1..108
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="TscAlu2"
repeat_region 1..108
            /note="scAlu"
            /rpt_family="Alu"

BASE COUNT 23 a 39 c 30 g 16 t
ORIGIN

Query Match 0.3%; Score 89; DB 11; Length 108;
Best Local Similarity 94.8%; Pred. No. 7.4e-08;
Matches 92; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9312 GCCTGTAATCCAGCACATTGGGAGCGCGAGCGGGCGGATCACGAGTTCAGGATGGA 9371
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 1 GCCTGTAATCCAGCACATTGGGAGCGCGAGCGGGCGGATCACGAGTTCAGGATGGA 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 9372 GACCATCTGCTTAAACACGATGAAACCCGCTCTCTAC 9408
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 61 GACCATCTGCTTAAACAGGTGAAACCCGCTCTCTAC 97
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 3
LOCUS HSLDLRN2 108 bp DNA PRI 20-MAY-1992
DEFINITION Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION X05250
VERSION X05250.1 GI:34337
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
Williamson,R. and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901
COMMENT See X05252 for deletion junction
FEATURES
    source
        Location/Qualifiers
            intron
                /db_xref="taxon:9606"
                /note="intron XIV fragment"
BASE COUNT 28 a 23 c 39 g 18 t
ORIGIN

Query Match 0.3%; Score 87.4; DB 10; Length 108;
Best Local Similarity 89.5%; Pred. No. 1.5e-07;
Matches 94; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 11837 CTCGGCTCAGCGCAACCTCCGCTCCAGGGTTCAGCAATTCCTCGCTCAGCTCCCC 11896
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 108 CTCGGCTCAGTCAACCTCTGCTCTCGGTTCAAGCAATTCCTCGCTCAGCTCCCG 49
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 11897 AGTAATCGGACTACTGGCAAGCGCCACGCTCGCTGAATTTT 11941
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 48 AGTAGCTGGGATTACAGCACCTGCCACACGCTGGCTAATTTT 4
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 4
LOCUS HSLDLRN2 108 bp DNA PRI 20-MAY-1992
DEFINITION Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION X05250
VERSION X05250.1 GI:34337
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
```

1 (bases 1 to 108)
 Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
 Williamson,R. and Humphries,S.
TITLE
 Unequal crossing-over between two alu-repetitive DNA sequences in
 the low-density-lipoprotein-receptor gene. A possible mechanism for
 the defect in a patient with familial hypercholesterolaemia
 Eur. J. Biochem. 164 (1), 77-81 (1987)
JOURNAL
 8/161901
MEDLINE
 See X05252 for deletion junction
COMMENT
 Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES
 source 1..108
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 intron 1..108
 /note="intron XIV fragment"
 BASE COUNT 28 a 23 c 39 g 18 t
 ORIGIN

Query Match 0.3%; Score 86.8; DB 10; Length 108;
 Best Local Similarity 88.7%; Pred. No. 2e-07;
 Matches 94; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 2629 AAAAAATTAGCGGGCGCTGTGGCGGGCGCCCTGTAGTCCACGCTACTTGGGAGGCTGAGGC 2888
 Db 3 AAAAAATTAGCGAGCGGTGGTGGCAGGTGCCCTGTATCCACGCTACTCTGGGAGGCTGAGGC 62
 QY 2889 AGGAGAATGGCATGAACCTGGGAGGCGCGAGCTTGCACTGAGCCGAG 2934
 Db 63 AGGAGATTGTTGACCCAGGAGCGAGGCTTGCACTGAGCCGAG 108

RESULT 5
 HUMALCE162/c HUMALCE162 107 bp ss-RNA PRI 15-APR-1994
LOCUS
 DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE162.
 ACCESSION M87924
 VERSION M87924.1 GI:174871
 KEYWORDS Alu repeat.
 SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
 1 (bases 1 to 107)
AUTHORS
 Slnnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE
 Alu RNA transcripts in human embryonal carcinoma cells. Model of
 post-transcriptional selection of master sequences
 J. Mol. Biol. (1992) In press
JOURNAL
 Location/Qualifiers
FEATURES
 source 1..107
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_line="NTera2D1"
 /dev_stage="embryo"
 /sex="male"
 /tissue_type="carcinoma"
 BASE COUNT 28 a 30 c 35 g 14 t
 ORIGIN

Query Match 0.3%; Score 85.4; DB 9; Length 107;
 Best Local Similarity 89.3%; Pred. No. 3.9e-07;
 Matches 92; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 10702 TTTTGTGAGACTGAGCCCTGCTCTGTCAACCGAGGTGGAGTGCAATGGCGGATCTCGGC 10761
 Db 107 TTTTGTGAGCGGAGTCTCGCTCTGTCCGCCAGGCTGGAGTGCACTGGCGGATCTCGGC 48

QY 10762 TCATGCAACCTCCGCTTCCAGGTTCAAGCGATTCTACTGCC 10804
 Db 47 TCATGCAAGCTCCGCCCTCCGGGGTTCACGCCATTCTCTCTGCC 5

QY	16540	CTGAGTGCAGTGGCAGACTTCGGCTCACTGCAACCTCTGTCTCCAGGGTTCAAGCGAT	16599
Db	103	CTGAGTGCAAATGGCAGACTTCGGCTCACTGCAACCTCCGGGTTCAAGCGAT	44
QY	16600	TCCTCCCTCACCCCTATGGAGTAGCTGGGATTACAGC	16638
Db	43	TCCTCCCTTAGCTTCCCGCTGTAGCTGGGATTACAGC	5
RESULT	8		
LOCUS	HSU67803/c		
DEFINITION	Human small cytoplasmic Alu transcript.		
ACCESSION	U67803		
VERSION	U67803.1	GI:2289917	
KEYWORDS	Alu.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Shalkh, T.H., Roy, A.M., Kim, J., Batzer, M.A. and Deininger, P.L.		
TITLE	CDNAs derived from primary and small cytoplasmic Alu (scAlu) transcripts		
JOURNAL	J. Mol. Biol.	271 (2), 222-234 (1997)	
MEDLINE	97415756		
REFERENCE	2 (bases 1 to 108)		
AUTHORS	Shalkh, T.H., Kim, J., Batzer, M.A. and Deininger, P.L.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The Children's Hospital of Philadelphia, 1004F Abramson Research Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA		
FEATURES	Location/Qualifiers		
source	1..108		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="TscAlu2"		
repeat_region	1..108		
	/note="scAlu"		
	/rpt_family="Alu"		
	/rpt_type="dispersed"		
BASE COUNT	23 a 39 c 30 g	16 t	
ORIGIN			
Query Match	0.3%;	Score 81;	DB 11;
Best Local Similarity	89.7%;	Pred. No. 2.9e-06;	Length 108;
Matches	87;	Conservative 0;	Mismatches 10;
		Indels 0;	Gaps 0;
QY	10871	GTAGACAGGGTTTTCACCGTGTGGCAGGATGTTCTCAATCTCTACCTCGTGATCC	10930
Db	97	GTAGACAGGGGTTTTCACCTGTGTACGAGGATGGTCTCGATCTCTGACCTCGTGATCC	38
QY	10931	GCCCGCTCTCTCTGCGAAAGTCTCGGATTACAGAC	10967
Db	37	GCCCGCTCGGCTCCAAAGTCTCGGATTACAGC	1
RESULT	9		
LOCUS	HSLDLR1		
DEFINITION	Human LDL-receptor mutated gene with intron 12 deletion junction.		
ACCESSION	X05249		
VERSION	X05249.1	GI:34335	
KEYWORDS	Alu repetitive sequence; low density lipoprotein receptor.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 108)		
	Horsthemke, B., Beisiegel, U., Dunning, A., Havinga, J.R., Williamson, R. and Humphries, S.		

TITLE	Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia					
JOURNAL	Eur. J. Biochem. 164 (1), 77-81 (1987)					
MEDLINE	87161901					
COMMENT	*source: hypercholesterol aemia See X05248 for corresponding normal gene sequence In the defective LDL-receptor gene the deletion occurred between two alu-repetitive sequences, that are in the same direction, the deletion eliminates exons 13 and 14 and changes the reading frame of the resulting spliced mRNA. Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.					
FEATURES	Location/Qualifiers					
source	1..108 /organism="Homo sapiens" /db_xref="taxon:9606" /cell_type="blood leukocytes from a patient with familial"					
misc_feature	1..108 /note="deletion junction region intron 12/ intron 15"					
BASE COUNT	20 a 40 c 20 g 28 t					
ORIGIN						
Query Match	0.3%; Score 80; DB 10; Length 108;					
Best Local Similarity	85.6%; Pred. No. 4.6e-06;					
Matches	89; Conservative 0; Mismatches 15; Indels 0; Gaps 0;					
QY 11838	TCGCCTCACCAGCAACCTCCGCTCCAGGGTTCACAGCAATTCTCCTCGCTCAGGCTCCCCA 11897 Db 2 TCGCCTCACCACACACTCTGCCTCCTGGGTTCAAACCATTTTCTCGCTCAGGCTCCCCA 61					
QY 11898	GTATTCGGGACTACTGCCAACGCCACACGCTCGTGCTAATTTT 11941 Db 62 GTACTGGGATTACAGCACCTCGCCACACGCTGGCTAATTT 105					
RESULT 10						
HSLDLRD2/c						
LOCUS	HSLDLRD2 108 bp DNA PRI 20-MAY-1992					
DEFINITION	Human LDL-receptor mutated gene with intron 14 deletion junction.					
ACCESSION	X05251					
VERSION	X05251.1 GI:34336					
KEYWORDS	Alu repetitive sequence; low density lipoprotein receptor.					
SOURCE	human.					
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 108) Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R., Williamson,R. and Humphries,S. unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia Eur. J. Biochem. 164 (1), 77-81 (1987)					
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
MEDLINE						
COMMENT	*source: hypercholesterol aemia See X05250 for corresponding normal gene sequence In the defective LDL-receptor gene the deletion occurred between two alu-repetitive sequences, that are in the same direction, the deletion eliminates exons 13 and 14 and changes the reading frame of the resulting spliced mRNA. Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.					
FEATURES	Location/Qualifiers					
source	1..108 /organism="Homo sapiens" /db_xref="taxon:9606" /cell_type="blood leukocytes from a patient with familial"					
intron	1..108 /note="intron XIV fragment"					
BASE COUNT	28 a 20 c 40 g 20 t					
ORIGIN						

Query Match	0.3%	Score 80;	DB 10;	Length 108;
Best Local Similarity	85.6%	Pred. No. 4.6e-06;		
Matches	89;	Conservative	0;	Mismatches 15; Indels 0; Gaps 0;
QY 11838	TCGGCTCAGCCCAACCTCCGGCTCCAGGGTTCAGAGCAATTCCTCGCTCAGCTCAGCCTCCGCCA 11857			
Db 107	TGGCTCACCACACACCTCTGCCTCTGGTTCACCAACATTTTCTGCTCAGCTCCCGA 48			
QY 11898	GTAAATGGGACTACTGGCAGCGCCACACCGCTGGCTAAATTTT 11941			
Db 47	GTAGCTGGGATTACAGCACCTGCACACCGCTGGCTAAATTTT 4			
RESULT 11				
LOCUS	HSLLDRD1	108 bp	DNA	PRI 20-MAY-1992
DEFINITION	Human LDL-receptor mutated gene with intron 12 deletion junction.			
ACCESSION	X05249			
VERSION	X05249.1	GI:34335		
KEYWORDS	Alu repetitive sequence; low density lipoprotein receptor.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;			
AUTHORS	Primates; Catarrhini; Homnidae; Homo.			
TITLE	1 (bases 1 to 108)			
JOURNAL	Horschman,B., Betsiegel,U., Dunning,A., Havinga,J.R.,			
MEDLINE	Williamson,R. and Humphries,S.			
COMMENT	Unequal crossing-over between two alu-repetitive DNA sequences in the low-density lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia			
FEATURES	Eur. J. Biochem. 164 (1), 77-81 (1987)			
LOCUS	87161901			
DEFINITION	*source: hypercholesterol aemia			
ACCESSION	See X05248 for corresponding normal gene sequence			
VERSION	In the defective LDL-receptor gene the deletion occurred between two			
KEYWORDS	alu-repetitive sequences, that are in the same direction, the			
SOURCE	deletion eliminates exons 13 and 14 and changes the reading frame			
ORGANISM	of the resulting spliced mRNA.			
DEFINITION	Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.			
ACCESSION	Location/Qualifiers			
VERSION	1..108			
DEFINITION	/organism="Homo sapiens"			
ACCESSION	/db_xref="taxon:9606"			
VERSION	/cell_type="blood leukocytes from a patient with familial"			
DEFINITION	misc_feature			
ACCESSION	1..108			
VERSION	/notes="deletion junction region intron 12/ intron 15"			
DEFINITION	BASE COUNT	20 a	40 c	20 g 28 t
ACCESSION	ORIGIN			
DEFINITION				
LOCUS				
DEFINITION	Query Match	0.3%	Score 79.4;	DB 10; Length 108;
ACCESSION	Best Local Similarity	84.8%	Pred. No. 6.1e-06;	
VERSION	Matches	89;	Conservative	0; Mismatches 16; Indels 0; Gaps 0;
KEYWORDS				
SOURCE				
ORGANISM				
DEFINITION	QY 2829	AAAAATTAGCGGCGTGTGGGGGGCGCTGTAGTCCACGCTACTTGGGAGCTGAGGC 2888		
ACCESSION	Db 106	AAAAATTAGCGGCGTGTGGGGGGCGCTGTAGTCCACGCTACTTGGGAGCTGAGGC 47		
VERSION	QY 2889	AGGAGATGCGATGAACCTGGGAGCGGAGCTTCGAGTGAGCCGA 2933		
KEYWORDS	Db 46	AGGAAATGTTTGAACCCAGGAGGAGGTTGTTGGTGGGCGA 2		
SOURCE				
ORGANISM				
DEFINITION	RESULT 12			
LOCUS	HSLLDRD2	108 bp	DNA	PRI 20-MAY-1992
DEFINITION	Human LDL-receptor mutated gene with intron 14 deletion junction.			
ACCESSION	X05251			
VERSION	X05251.1	GI:34336		
KEYWORDS	Alu repetitive sequence; low density lipoprotein receptor.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;			

Result No.	Score	Query		Length	DB	ID	Description
		Match					
C 1	71.6	0.2	108	1	T26828	Human gene signatu	
C 2	67.2	0.2	100	1	T24892	Human gene signatu	
C 3	66.8	0.2	108	1	X12095	Human biallelic po	
C 4	63.8	0.2	86	1	Y41231	Mouse embryonic ce	
C 5	64	0.2	100	1	T24892	Human gene signatu	
C 6	62.8	0.2	103	1	T20927	Human gene signatu	
C 7	62.8	0.2	108	1	X12095	Human biallelic po	
C 8	61	0.2	108	1	T25009	Human gene signatu	
C 9	60	0.2	93	1	T25688	Human gene signatu	
C 10	60.2	0.2	108	1	T25009	Human gene signatu	
C 11	59.2	0.2	100	1	X12087	Human biallelic po	
C 12	59.2	0.2	100	1	X12085	Human biallelic po	
C 13	58.6	0.2	110	1	T26288	Human gene signatu	
C 14	58	0.2	100	1	X12086	Human biallelic po	
C 15	57.4	0.2	110	1	T25260	Human gene signatu	
C 16	57	0.2	108	1	T26828	Human gene signatu	
C 17	56.4	0.2	99	1	T20931	Human gene signatu	
C 18	56.2	0.2	103	1	T26213	Human gene signatu	
C 19	55	0.2	100	1	X12087	Human biallelic po	
C 20	55	0.2	100	1	X12085	Human biallelic po	
C 21	55	0.2	100	1	X12086	Human biallelic po	
C 22	54.6	0.2	69	1	Q29016	Probe to internal	
C 23	54.6	0.2	91	1	T25854	Human gene signatu	
C 24	54.8	0.2	97	1	T26728	Human gene signatu	
C 25	54.4	0.2	109	1	T23895	Human gene signatu	
C 26	53.4	0.2	97	1	T26728	Human gene signatu	
C 27	53.6	0.2	100	1	Q76490	Human genome fragm	
C 28	53.2	0.2	82	1	T25468	Human gene signatu	
C 29	53.2	0.2	110	1	T26288	Human gene signatu	
C 30	51.6	0.2	89	1	T23513	Human gene signatu	
C 31	51.4	0.2	91	1	T25854	Human gene signatu	
C 32	51.4	0.2	103	1	T26213	Human gene signatu	
C 33	50.4	0.2	62	1	T25689	Human gene signatu	
C 34	50.4	0.2	102	1	T20743	Human gene signatu	


```
RESULT 10
T25009
ID T25009 standard; cDNA to mRNA; 108 BP.
AC
DT 07-NOV-1996 (first entry)
DE Human gene signature HUMGS07131.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-AL.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues.
PS Claim 1; Page 1748; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 34 A; 31 C; 26 G; 15 T;

Query Match 0.2%; Score 60.2; DB 1; Length 108;
Best Local Similarity 72.0%; Pred. No. 0.12;
Matches 77; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 9486 ATGGCGTGAACGGGGAGCGGAGCTTGCAGTGCAGGAGCGGAGTGCAGGAGCGGAGTGCAG 9545
DB 2 ATCGCCTGAGCCCATGAGGCGCAAGGCTGCAGTGCAGGAGCGGAGTGCAGGAGCGGAGTGCAG 61
QY 9546 CCTGGGTGCAGACGAGAGCTCCGTCTCAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9592
DB 62 CCTGAGTGCAGACGAGACGAGCCCTGTTGAAAAACACACACACACACACACACACACAC 108

RESULT 11
X12087/c
ID X12087 standard; DNA; 100 BP.
AC X12087;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment EST98276a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.

PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match 0.2%; Score 59.2; DB 1; Length 100;
Best Local Similarity 74.5%; Pred. No. 0.17;
Matches 73; Conservative 1; Mismatches 24; Indels 0; Gaps 0;

QY 8434 TGCAGCTCAGCCCTGTAATACACGACATTTTGGGAGGCCAAGGTGGGAGGATCACTTGCAGC 8493
DB 100 TGTGACTCACACCTATTAATCTTGCACCTTTAGGAGGCTTAGGAAGGAGGATGTTTGA 41
QY 8494 CCAGGAGCTCAAGACACGAGCTCTGGGCAACTTAGTGAGAC 8531
DB 40 CCAGGAGCTCAAGACCAKCTCTGGGAAACATAGCAAGAC 3

RESULT 12
X12085/c
ID X12085 standard; DNA; 100 BP.
AC X12085;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment EST98276c.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.

PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 218; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
```

CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 22 A; 25 C; 22 G; 30 T;

Query Match 0.2%; Score 59.2; DB 1; Length 100;
Best Local Similarity 74.5%; Pred. No. 0.17;
Matches 73; Conservative 1; Mismatches 24; Indels 0; Gaps 0;

QY 8434 TGCAGCTCAGGCTGTATACCAAGCATTTGGGAGGCCAAGGTGGGAGGATCATTGAGC 8493
II IIIII III IIII III IIII IIII IIII IIII IIII IIII IIII
DB 100 TGTGACTCACCTATATCTCTGGCACTTTGGAGGCTTAGGAGGAGGATTTGTTGAA 41

QY 8494 CCAGGAGCTCAAGACCACTCTGGGCACTTAGTGAGAC 8531
IIIIII IIIII IIII IIII IIII IIII IIII IIII IIII IIII
DB 40 CCAGGAGCTCAAGACCACTCTGGGAAACATAGCAAGAC 3

RESULT 13
T26288
ID T26288 standard; cDNA to mRNA; 110 BP.
AC T26288;
DE Human gene signature HUMG08527.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-Al.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATSU) MATSUBARA K.
PA (OKUBU) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 2048; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 110 BP; 22 A; 37 C; 28 G; 17 T;

Query Match 0.2%; Score 58.6; DB 1; Length 110;
Best Local Similarity 85.3%; Pred. No. 0.22;
Matches 64; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 12008 GATCCACCGGCTCGGCTCCCAAGTCTGGGATTCAGAGTATGAGCACTGGGCCCGG 12087
IIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII
DB 1 GATCCGCGGCTCGAGCTCCCAAGTCTGGGATTCAGAGTATGAGCACTGGCACCGG 60

QY 12068 CCACATTTCTAAAT 12082

DB 61 CCCCATTCCTCACTT 75
II IIII IIII IIII

RESULT 14
X12086/c
ID X12086 standard; DNA; 100 BP.
AC X12086;
DE Human biallelic polymorphic DNA fragment EST98276b.
DE Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match 0.2%; Score 58; DB 1; Length 100;
Best Local Similarity 74.5%; Pred. No. 0.27;
Matches 73; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 8434 TGCAGCTCAGGCTGTATACCAAGCATTTGGGAGGCCAAGGTGGGAGGATCATTGAGC 8493
II IIIII IIII IIII IIII IIII IIII IIII IIII IIII IIII
DB 100 TGTGACTCACCTATATCTCTGGCACTTTAGGAGGCTTAGGAGGAGGATTTGTTGAA 41

QY 8494 CCAGGAGCTCAAGACCACTCTGGGCACTTAGTGAGAC 8531
IIIIII IIIII IIII IIII IIII IIII IIII IIII IIII IIII
DB 40 CCAGGAGCTCAAGACCACTCTGGGAAACATAGCAAGAC 3

RESULT 15
T25260/c
ID T25260 standard; DNA; 110 BP.
AC T25260;
DE Human gene signature HUMG07421.
DE Human gene signature HUMG07421.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-Al.
PD 01-JUN-1995.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2000, 12:40:45 ; Search time 8475.2 Seconds
(without alignments)
13869.588 Million cell updates/sec

Title: US-08-852-495C-2_COPY_84000_113000
Perfect score: 29001
Sequence: 1 TCAAACTCCTGACCTCATGA.....ATAAATATCTTAAATAACC 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result % Query

No.	Score	Match	Length	DB	ID	Description
C 1	96.2	0.3	109	30	AA243009	zr25h02.s
C 2	94.8	0.3	106	37	AA703692	ag81a10.r
C 3	93.8	0.3	105	61	AI832832	at72909.x
C 4	93.2	0.3	106	38	AA812141	ob48h02.s
C 5	91.8	0.3	103	84	B48914	RPC111-4A12
C 6	91.4	0.3	101	39	AA835205	ak64h01.s
C 7	91.6	0.3	109	24	N25299	yw52c09.s1
C 8	91.4	0.3	109	30	AA243009	zr25h02.s
C 9	91.6	0.3	110	39	AA897366	am06h02.s
C 10	91.2	0.3	107	39	AA828124	od71a07.s
C 11	89.6	0.3	97	25	N49638	yv25e09.r1
C 12	88.2	0.3	103	94	AQ028649	CIT-HSP-2
C 13	88.2	0.3	101	94	AQ076649	CIT-HSP-2
C 14	88.2	0.3	102	30	AA226656	nc19f09.s
C 15	88.4	0.3	106	105	AQ264176	CITBI-E1-
C 16	88	0.3	107	33	AA385808	EST99495
C 17	87.6	0.3	103	38	AA807640	nx08b05.s
C 18	87.6	0.3	110	79	AW250394	2822460.3
C 19	86.8	0.3	106	37	AA703692	ag81a10.r
C 20	85.8	0.3	106	108	AQ5444957	CITBI-E1-
C 21	85.8	0.3	107	24	H67040	yu68c01.r1
C 22	86	0.3	110	64	AW083640	xc49f02.x
C 23	85.4	0.3	103	84	B48914	RPC111-4A12
C 24	85	0.3	109	84	B17434	345K2.TVB C
C 25	85	0.3	110	94	AQ003188	RPC111-1D
C 26	84.4	0.3	103	108	AQ582186	RPC111-4
C 27	83.8	0.3	103	30	AA228795	nc14e07.s
C 28	83.8	0.3	105	109	AQ637292	RPC1-11-4
C 29	84	0.3	110	33	AA442529	zv68b02.r
C 30	83.6	0.3	106	34	AA516339	ng71g02.s
C 31	83.6	0.3	107	35	AA565533	nk42b11.s
C 32	83.2	0.3	107	24	N23686	yw46a02.s1
C 33	83.4	0.3	109	94	AQ028426	CIT-HSP-2
C 34	83.4	0.3	110	94	AQ003188	RPC111-1D
C 35	82.8	0.3	102	36	AA654562	nt75f10.s
C 36	82.8	0.3	105	74	AW196212	xm06e06.x
C 37	83	0.3	107	62	AI933497	wm74d02.x
C 38	82.4	0.3	93	43	AI168167	oc09e10.x
C 39	82.6	0.3	108	32	AA370029	EST81584
C 40	82	0.3	100	35	AA564832	nj22a06.s
C 41	82.2	0.3	103	94	AQ028649	CIT-HSP-2
C 42	82.2	0.3	103	108	AQ584425	RPC1-11-4
C 43	81.8	0.3	101	35	AA583697	nn58f10.s
C 44	82	0.3	106	38	AA812141	ob48h02.s
C 45	82	0.3	106	94	AQ062963	CIT-HSP-2

ALIGNMENTS

RESULT 1
AA243009/c
LOCUS AA243009 109 bp mRNA EST 11-MAR-1998
DEFINITION zr25h02.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
cDNA clone IMAGE:664467 3' similar to contains Alu repetitive
element; contains element LTR1 repetitive element ;, mRNA sequence.

ACCESSION AA243009.1 GI:1873869
VERSION EST.
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 109)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Dec 3, 1996 this sequence version replaced gi:1126869.

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1127 Std Error: 0.00
Seq primer: -4ml3 fwd. ST from Amersham
High quality sequence stop: 102.

FEATURES
source

Location/Qualifiers
1..109
/organism="Homo sapiens"
/db_xref="GDB:5426481"
/db_xref="taxon:9606"
/clone="IMAGE:664467"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/tissue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: brain; Vector: pBluescript SK-; Site: 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Uninduced, exponentially growing neuroepithelial
cells (Ntera-2/ci.D1). Average insert size: 1.0 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3' "
BASE COUNT 19 a 30 c 30 g 30 t
ORIGIN

Query Match 0.3%; Score 96.2; DB 30; Length 109;
Best Local Similarity 92.7%; Pred. No. 0.076; 8; Indels 0; Gaps 0;
Matches 101; Conservative 0; Mismatches 0;
Qy 9309 CACGCTTAATCCAGCACCTTTGGAGCGGAGCGGCGGATCAGGATCAGAGAT 9368
Db 109 CACGCTTAATCCAGCACCTTTGGAGCGGAGCGGCGGATCAGGATCAGAGAT 50
Qy 9369 GGAGACCATCTCTGTACACAGATGAACCCCGTCTCTACTAAAAATAC 9417
Db 49 CAAGACCATCTCTGTACACAGGTAACCCCGTCTCTACTAAAAATAC 1

RESULT 2
AA703692/c AA703692 106 bp mRNA EST 24-DEC-1997
LOCUS AA703692.1 GI:2713610
DEFINITION ag81a10.r1 Stratagene hnt neuron (#937233) Homo sapiens cDNA clone
IMAGE:1140858 5' similar to contains Alu repetitive element; , mRNA
sequence.

ACCESSION AA703692
VERSION AA703692.1 GI:2713610
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 106)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1397630.
Contact: Wilson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 53.

FEATURES source

1. .106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1140858"
/clone_lib="Stratagene hMT neuron (#937233)"
/dev_stage="hMT neurons"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Differntiated, post mitotic hMT neurons. Average insert size: 1.5 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGAG 3' ~3' adaptor sequence: 5' CTCGAGTCTTTTCTTTT 3' "

BASE COUNT 19 a 29 c 29 g 29 t
ORIGIN

Query Match 0.3%; Score 94.8; DB 37; Length 106;
Best Local Similarity 93.4%; Pred. No. 0.11; 7; Indels 0; Gaps 0;
Matches 99; Conservative 0; Mismatches 0;

QY 9309 CACGCTGTAAATCCAGCACTTTGGAGCGCGGCGGATCAGGAGTCAAGGAT 9368
Db 106 CACGCTGTAAATCCAGCACTTTGGAGCGCGGCGGATCAGGAGTCAAGGAT 47
QY 9369 GGAGACCATCTCGTTAAACAGATGAACCCCGTCTCTACTAAAAA 9414
Db 46 CGAGACCATCTCGTTAAACAGATGAACCCCGTCTCTACTAAAAA 1

RESULT 3

AI832832 105 bp mRNA EST 13-JUL-1999
LOCUS at72g09.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone
DEFINITION IMAGE:2377600 3' similar to contains Alu repetitive
element; contains element MER22 repetitive element ;, mRNA sequence.
ACCESSION AI832832
VERSION AI832832.1 GI:5454812
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 105)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Dec 20, 1995 this sequence version replaced gi:1133644.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco.

FEATURES source

1. .105
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2377600"
/clone_lib="Barstead colon HPLRB7"
/sex="male"
/dev_stage="adult, age 25"
/lab_host="DH10B (phage resistant)"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a

modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dt) primer [5'
TGCTACGATCTGAAGTGGAGCGGCCCTTTTCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[5' AATTCACCTAGTAAT 3' and 5' ATTACTAGT 3'], digested
with Not I and cloned into the Not I and Eco RI sites of
the modified pT73 vector. Library constructed by Bob
Barstead."

BASE COUNT 17 a 35 c 27 g 26 t
ORIGIN

Query Match 0.3%; Score 93.8; DB 61; Length 105;
Best Local Similarity 93.3%; Pred. No. 0.14; 7; Indels 0; Gaps 0;
Matches 98; Conservative 0; Mismatches 0;

QY 11788 GAGACGGAGTTTCACACTTGTCCCGCAGGCTGGAGTGAATGTGGATCTCGGCTCACC 11847
Db 1 GAGACAGAGTTTCGCTCTCTTGTCCCGCAGGCTGGAGTGAATGTGGATCTCGGCTCACC 60
QY 11848 GCAACCTCCGCTCCAGGTTCAAGCAATTCCTCCTCAGCCT 11892
Db 61 GCAACCTCCAGCTCCCGGGTCAAGCGATTCTCCTCGCTCAGCCT 105

RESULT 4

AA812141/c 106 bp mRNA EST 19-FEB-1998
LOCUS ob48h02.sl NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1334643 3'
DEFINITION similar to contains Alu repetitive element; , mRNA sequence.
ACCESSION AA812141
VERSION AA812141.1 GI:2881752
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 106)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1402063.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html

Insert Length: 1450 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 60.

FEATURES source

1. .106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1334643"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was

primed with a Not I - oligo(dT) primer
[5'-TGTTCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTTTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT773 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 16 a 31 c 24 g 35 t
ORIGIN

Query Match 0.3%; Score 93.2; DB 38; Length 106;
Best Local Similarity 92.5%; Pred. No. 0.17;
Matches 98; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 9481 GGAGATGCGTGAACGGGAGCGGAGCTTGCAGTCAGCGGAGATCGCGCCAGCGCAC 9540

Db 106 GGAGATGCGTGAACCGTGGGAGGTGGAGCTTGCAGTCAGCGGAGATCACACCTGCAC 47

QY 9541 TCCAGCGCTGGTGCACAGCGAGACTCCGCTCTCAAAAAA 9586

Db 46 TCCAGCGCTGGTGCACAGCGAGACTCCATCTCAAAAAA 1

RESULT 5

LOCUS B48914

DEFINITION B48914 103 bp DNA GSS 08-APR-1999
RPC111-4A12.TP RPC1-11 Homo sapiens genomic clone RPC1-11-4A12,
genomic survey sequence.

ACCESSION B48914

VERSION B48914.1 GI:2601151

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venter,J.C.

TITLE Use of BAC End Sequences for Sequence-Ready Map Building

JOURNAL Unpublished (1997)

COMMENT Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are derived from the human BAC library RPC1-11. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from

Research Genetics (inforesgen.com). BAC end search page:

http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html

Seq primer: SP6

Class: BAC ends.

FEATURES

source

1..103

/organism="Homo sapiens"

/db_xref="GDB:7501163"

/db_xref="taxon:9606"

/clone="RPC1-11-4A12"

/clone_lib="RPC1-11"

/sex="Male"

/cell_type="Lymphocytes"

/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;

RPC111 Human Male BAC Library"

30 a 28 c 30 g 15 t

BASE COUNT

ORIGIN

Query Match 0.3%; Score 91.8; DB 84; Length 103;

Best Local Similarity 93.2%; Pred. No. 0.24;
Matches 96; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 9316 GTAATCCAGCACTTTGGAGGCGGAGCGGATCACGAGTTCAGGATGAGAC 9375

Db 1 GTAAGCCAGCACTTTGGAGGCGGAGCGGATCACGAGTTCAGGATGAGAC 60

QY 9376 ATCTCTGCTTAACACGATGAAACCCCGTCTCTACTAAAAATACA 9418

Db 61 ATCCGGCTAAACGGTGAACCCCGTCTCTACTAAAAATACA 103

RESULT 6

LOCUS AA835205/c

DEFINITION at64h01.ai Barstead pancreas HPLRB1 Homo sapiens cDNA clone

IMAGE:1412689 3' similar to contains Alu repetitive

element; contains element KER repetitive element ; , mRNA sequence.

ACCESSION AA835205

VERSION AA835205.1 GI:2908933

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 101)

Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,

Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,

Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,

Theising,B., White,Y., Wyllie,F., Waterston,R. and Wilson,R.

WashU-NCI human EST Project

Unpublished (1997)

On Nov 29, 1993 this sequence version replaced gi:636191.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (infoimage.lnl.gov) for further information.

Seq primer: -40m13 fwd. EF from Amersham.

FEATURES

source

1..101

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1412689"

/clone_lib="Barstead pancreas HPLRB1"

/sex="female"

/dev_stage="adult, 34 years"

/lab_host="DH10B"

/note="Organ: pancreas; Vector: pT7T3D-Pac (Pharmacia)

with a modified polylinker; Site_1: EcoRI; Site_2: NotI;

1st strand cDNA was primed with a Not I - oligo(dT) primer

[5'

TGTTACGAATCGAAGTGGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT

3']; double-stranded cDNA was ligated to Eco RI adaptors

[AATTCGATCCTTG], digested with Not I and cloned into the

Not I and Eco RI sites of the modified pT7T3 vector.

Library constructed by Bob Barstead."

14 a 36 c 27 g 24 t

BASE COUNT

ORIGIN

Query Match 0.3%; Score 91.4; DB 39; Length 101;

Best Local Similarity 94.1%; Pred. No. 0.27; Indels 0; Gaps 0;

Matches 95; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 9474 TGAGGAGAGAGTGGCTGAACGGGAGCGGAGCTTGCAGTGCAGCGAGATCGGCC 9533

Db 101 TGAGGAGAGAGTGGCTGAACCGGGAGCGGAGCTTGCAGTGCAGCGAGATCAAGCC 42

QY 9534 ACGGCACCTCCAGCCTGGGTGCAGAGCGAGAGACTCCGCTCTCA 9574


```

ACCESSION AA897366
VERSION AA897366.1 GI:3033986
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 110)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2150764.
Contact: Robert Strausberg, Ph.D.
Tel.: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 834 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 63.
FEATURES
Location/Qualifiers
1..110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1466067"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI-CGAP-GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 22 a 27 c 29 g 32 t
ORIGIN
Query Match 0.3%; Score 91.6; DB 39; Length 110;
Best Local Similarity 91.5%; Pred. No. 0.24;
Matches 97; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 9309 CACGGCTGTATCCAGCACCTTTGGAGCGCCGAGCGCGGATCACGAGGTCCAGAGAT 9368
|||||
Db 110 CACGCTATATCCAGCACCTTTGGAGCGCCGAGGTCCAGATCATGAGTCCAGAGAT 51

QY 9369 GGAGACCATCTCTGTTAACACGATGAACCCCGTCTCTACTATAAAA 9414
|||||
Db 50 TGAGACCATCTCTGTTAACACGCGTGAACCCCATCTCTACTATAAAA 5

RESULT 10
AA828124 107 bp mRNA EST 20-FEB-1998
LOCUS od71a07.s1 NCI-CGAP_Ov2 Homo sapiens cDNA clone IMAGE:1373364
DEFINITION similar to contains Alu repetitive element;contains element MER22
repetitive element ;, mRNA sequence.
ACCESSION AA828124
VERSION AA828124.1 GI:2900487
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 107)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2150764.
Contact: Robert Strausberg, Ph.D.
Tel.: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 834 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 63.
FEATURES
Location/Qualifiers
1..110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1466067"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI-CGAP-GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 22 a 27 c 29 g 32 t
ORIGIN
Query Match 0.3%; Score 91.6; DB 39; Length 110;
Best Local Similarity 91.5%; Pred. No. 0.24;
Matches 97; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 9309 CACGGCTGTATCCAGCACCTTTGGAGCGCCGAGCGCGGATCACGAGGTCCAGAGAT 9368
|||||
Db 110 CACGCTATATCCAGCACCTTTGGAGCGCCGAGGTCCAGATCATGAGTCCAGAGAT 51

QY 9369 GGAGACCATCTCTGTTAACACGATGAACCCCGTCTCTACTATAAAA 9414
|||||
Db 50 TGAGACCATCTCTGTTAACACGCGTGAACCCCATCTCTACTATAAAA 5

RESULT 11
AA828124 107 bp mRNA EST 20-FEB-1998
LOCUS od71a07.s1 NCI-CGAP_Ov2 Homo sapiens cDNA clone IMAGE:1373364
DEFINITION similar to contains Alu repetitive element;contains element MER22
repetitive element ;, mRNA sequence.
ACCESSION AA828124
VERSION AA828124.1 GI:2900487
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 107)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 17, 1998 this sequence version replaced gi:1899815.
Contact: Robert Strausberg, Ph.D.
Tel.: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 93.
FEATURES
Location/Qualifiers
1..107
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1373364"
/clone_lib="NCI-CGAP_Ov2"
/sex="female"
/tissue_type="ovary"
/lab_host="DH10B"
/note="Vector: pAMP10; mRNA made from invasive ovarian
tumor, cDNA made by oligo-dT priming. Non-directionally
cloned. Size-selected on agarose gel, average insert size
600 bp. Reference: Krizman et al. (1996) Cancer Research
56:5380-5383."
BASE COUNT 30 a 23 c 38 g 16 t
ORIGIN
Query Match 0.3%; Score 91.2; DB 39; Length 107;
Best Local Similarity 92.3%; Pred. No. 0.27;
Matches 96; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2814 TCTACTAAATACAAAAATAGCGCGCTGTGGCGGCGCTGTAGTCCAGCTAC 2873
|||||
Db 4 TCGACTAAATACAAAAATAGCGCGCTGTGGCGGCGCTGTAGTCCAGCTAC 63

QY 2874 TTGGGAGCTGAGGAGAGATGGCATGAACCTGGAGCGGA 2917
|||||
Db 64 TTGGGAGCTGAGGAGAGATGGCGTGAACCCGGAGCGGA 107

RESULT 11
AA828124 107 bp mRNA EST 14-FEB-1996
LOCUS YV25e09.r1 Soares fetal liver spleen 1NPLS Homo sapiens cDNA clone
DEFINITION IMAGE:243784 5' similar to gb:X57138_rnal HISTONE H2B.2 (HUMAN);,
mRNA sequence.
ACCESSION N49638
VERSION N49638.1 GI:1190804
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 97)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT On Apr 14, 1993 this sequence version replaced gi:693230.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

```


Search completed: June 18, 2000, 22:02:37
Job time: 379524 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2000, 17:46:45 ; Search time 373.09 Seconds
(without alignments)
10103.987 Million cell updates/sec

Title: US-08-852-495C-2_COPY_84000_113000
Perfect score: 29001
Sequence: 1 TCAAACTCCTGACCTCATGA.....ATAAATATCTTTAAATACC 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2.6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2.6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2.6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn2.6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn2.6/ptodata/1/ina/6_COMB.seq:*
6: /cgn2.6/ptodata/1/ina/PTUS_COMB.seq:*
7: /cgn2.6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	80	0.3	105	4	US-08-481-658B-65
2	80	0.3	105	4	US-08-477-504A-65
3	80	0.3	105	4	US-08-486-756A-65
4	80	0.3	105	4	US-08-485-862B-65
5	80	0.3	105	5	US-08-787-739-65
6	76.2	0.3	105	4	US-08-481-658B-65
7	76.2	0.3	105	4	US-08-477-504A-65
8	76.2	0.3	105	4	US-08-486-756A-65
9	76.2	0.3	105	4	US-08-485-862B-65
10	76.2	0.3	105	5	US-08-787-739-65
11	67.6	0.2	84	3	US-08-454-557C-91
12	67.6	0.2	84	4	US-08-340-426B-91
13	67.6	0.2	84	4	US-08-450-673C-91
14	67.6	0.2	84	6	PCT-US95-17111A-91
15	62.8	0.2	76	3	US-08-454-557C-69
16	62.8	0.2	76	4	US-08-340-426B-69
17	62.8	0.2	76	4	US-08-450-673C-69
18	62.8	0.2	76	6	PCT-US95-17111A-69
19	59.2	0.2	85	3	US-08-454-557C-92
20	59.2	0.2	85	4	US-08-340-426B-92
21	59.2	0.2	85	4	US-08-450-673C-92
22	59.2	0.2	85	6	PCT-US95-17111A-92
23	58.4	0.2	60	3	US-08-454-557C-57
24	58.4	0.2	60	4	US-08-340-426B-57
25	58.4	0.2	60	4	US-08-450-673C-57
26	58.4	0.2	60	6	PCT-US95-17111A-57
27	57.4	0.2	84	3	US-08-454-557C-91

28	57.4	0.2	84	4	US-08-340-426D-91	Sequence 91, Appl
29	57.4	0.2	84	4	US-08-450-673C-91	Sequence 91, Appl
30	57.4	0.2	84	6	PCT-US95-17111A-91	Sequence 91, Appl
31	55	0.2	85	3	US-08-454-557C-92	Sequence 92, Appl
32	55	0.2	85	4	US-08-340-426B-92	Sequence 92, Appl
33	55	0.2	85	4	US-08-450-673C-92	Sequence 92, Appl
34	55	0.2	85	6	PCT-US95-17111A-92	Sequence 92, Appl
35	52	0.2	60	3	US-08-454-557C-60	Sequence 60, Appl
36	52	0.2	60	4	US-08-340-426B-60	Sequence 60, Appl
37	52	0.2	60	4	US-08-450-673C-60	Sequence 60, Appl
38	52	0.2	60	6	PCT-US95-17111A-60	Sequence 60, Appl
39	52.2	0.2	94	5	US-08-750-064-6	Sequence 6, Appl
40	50.8	0.2	78	3	US-08-454-557C-70	Sequence 70, Appl
41	50.8	0.2	78	4	US-08-340-426D-70	Sequence 70, Appl
42	50.8	0.2	78	4	US-08-450-673C-70	Sequence 70, Appl
43	50.8	0.2	78	6	PCT-US95-17111A-70	Sequence 70, Appl
44	50.4	0.2	60	5	US-08-545-860D-71	Sequence 71, Appl
45	50.4	0.2	60	6	PCT-US94-04496-71	Sequence 71, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-65
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-481-658B-65

Query Match 0.3%; Score 80; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. No. 9.5e-10;

Matches 89; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 10857 TTTTGTATTTTGTAGACAGGTTTCACCGTGTGGCCAGGATGTTCTCAATCTCC 10916

Db 2 TTTTACATCTTTAGTAGACAGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

QY 10917 TTACCTCGTGATCCGCCCTCGCTGCTGCCAAAGTGCTCGGAT 10960

Db 62 TGACCTTGATGCCACAGCCTCGGCTCCCAAAAGTGCTGGAT 105

RESULT 2

US-08-477-504A-65

; Sequence 65, Application US/08477504A

; Patent No. 5972353

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,504A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3D

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-477-504A-65

Query Match 0.3%; Score 80; DB 4; Length 105;

Best Local Similarity 85.6%; Pred. No. 9.5e-10;

Matches 89; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 10857 TTTTGTATTTTGTAGACAGGTTTCACCGTGTGGCCAGGATGTTCTCAATCTCC 10916

Db 2 TTTTACATCTTTAGTAGACAGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

QY 10917 TTACCTCGTGATCCGCCCTCGCTGCTGCCAAAGTGCTCGGAT 10960

Db 62 TGACCTTGATGCCACAGCCTCGGCTCCCAAAAGTGCTGGAT 105

RESULT 3

US-08-486-756A-65

; Sequence 65, Application US/08486756A

; Patent No. 5981711

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,756A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-486-756A-65

Query Match 0.3%; Score 80; DB 4; Length 105;

Best Local Similarity 85.6%; Pred. No. 9.5e-10;

Matches 89; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 10857 TTTTGTATTTTGTAGACAGGTTTCACCGTGTGGCCAGGATGTTCTCAATCTCC 10916

Db 2 TTTTACATCTTTAGTAGACAGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

QY 10917 TTACCTCGTGATCCGCCCTCGCTGCTGCCAAAGTGCTCGGAT 10960

Db 62 TGACCTTGATGCCACAGCCTCGGCTCCCAAAAGTGCTGGAT 105

RESULT 4

US-08-485-862B-65

; Sequence 65, Application US/08485862B

; Patent No. 5989838

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

```

CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,862B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-485-862B-65

Query Match 0.3%; Score 80; DB 4; Length 1
Best Local Similarity 85.6%; Pred. No. 9.5e-10;
Matches 89; Conservative 0; Mismatches 15; Indel

QY 10857 TTTTGTGATTTTAGTAGACAGAGGTTTCACCGTGTGGCCACGAGTGT
DB 2 TTTTTCATCTTTAGTAGACAGAGGTTTTTCCACATATTGGCCAGGCTGCT
QY 10917 TTACCTCGTGATCCGCCGCGCTGCTCCTCCCAAAGTCTCGGAT 10960
DB 62 TGACCTTGATGCCACAGCCTCGGCCTCCCAAAGTCTGGGAT 105

RESULT 5
US-08-787-739-65
Sequence 65, Application US/08787739
Patent No. 6027887
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 96
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 369 Pine Street, Suite 610
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:

```

```

/ APPLICATION NUMBER: US/08/787,739
/ FILING DATE: 24-JAN-1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/485,049
/ FILING DATE: 07-JUN-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/486,756
/ FILING DATE: 07-JUN-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/477,504
/ FILING DATE: 07-JUN-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/481,658
/ FILING DATE: 07-JUN-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/485,862
/ FILING DATE: 07-JUN-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/485,863
/ FILING DATE: 07-JUN-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/487,077
/ FILING DATE: 07-JUN-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Lauder, Leona L.
/ REGISTRATION NUMBER: 30,863
/ REFERENCE/DOCKET NUMBER: D-0021.4
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-981-2034
/ TELEFAX: 415-981-0332
/ INFORMATION FOR SEQ ID NO: 65:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 105 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/
/ US-08-787-739-65
/
/ Query Match 0.3%; Score 80;
/ Best Local Similarity 85.6%; Pred. No. 9
/ Matches 89; Conservative 0; Mismatchc
/
QY 10857 TTTTGTATTTTAGTAGACAGGCTTCACCGT
Db 2 TTTTACATCTTTAGTAGACAGGTTTTCACCAT
QY 10917 TTACCTCGTGATCGCGCGCTCGTCTGCGAAAG
Db 62 TGACCTTGTGATCCACGAGCTCGGCTCCCAAG
/
RESULT 6
/ US-08-481-658B-65/c
/ Sequence 65, Application US/08481658B
/ Patent No. 5955075
/ GENERAL INFORMATION:
/ APPLICANT: zavada, Jan
/ APPLICANT: Pastorekova, Silvia
/ APPLICANT: Pastorek, Jaromir
/ TITLE OF INVENTION: MN Gene and Protein
/ NUMBER OF SEQUENCES: 86
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Leona L. Lauder
/ STREET: 6 Mariposa Court
/ CITY: Tiburon
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94920
/ COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.3%; Score 76.2; DB 4; Length 105;
Best Local Similarity 82.9%; Pred. No. 7.3e-09;
Matches 87; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 9319 ATCCGAGCAGCTTTGGAGGCGGCGGATCAGGAGTGCAGGAGATGGAGACCATC 9378
|||||
DB 105 ATCCGAGCAGCTTTGGAGGCGGCGGATCAGGAGTGCAGGAGTGCAGGAGATGGAGACCATC 9378
|||||

QY 9379 CTGCTTAACACGATGAACCCGCTCTCTACTAAATAACAAATA 9423
|||||
DB 45 CTGGCCAATATGTTGAACCCGCTCTCTACTAAAGATGTAAAAA 1

RESULT 7
US-08-477-504A-65/c
Sequence 65, Application US/08477504A
Patent No. 5972353
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-477-504A-65

Query Match 0.3%; Score 76.2; DB 4; Length 105;
Best Local Similarity 82.9%; Pred. No. 7.3e-09;
Matches 87; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 9319 ATCCGAGCAGCTTTGGAGGCGGCGGATCAGGAGTGCAGGAGATGGAGACCATC 9378
|||||
DB 105 ATCCGAGCAGCTTTGGAGGCGGCGGATCAGGAGTGCAGGAGTGCAGGAGATGGAGACCATC 9378
|||||

QY 9379 CTGCTTAACACGATGAACCCGCTCTCTACTAAATAACAAATA 9423
|||||
DB 45 CTGGCCAATATGTTGAACCCGCTCTCTACTAAAGATGTAAAAA 1

RESULT 8
US-08-486-756A-65/c
Sequence 65, Application US/08486756A
Patent No. 5981711
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,756A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-486-756A-65

Query Match 0.3%; Score 76.2; DB 4; Length 105;
Best Local Similarity 82.9%; Pred. No. 7.3e-09;
Matches 87; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 9319 ATCCACGACTTTGGGAGCGGCGGCGGATCAGGATCAGGAGTGAGGACCATC 9378
Db 105 ATCCACGACTTTGGGAGCGGCGGCGGATCAGGATCAGGAGTGAGGACCATC 46
QY 9379 CTGCTTAAACAGATGAACCCCGTCTCTACTATAAAATACAAATA 9423
Db 45 CTGCCCAATATGGTGAACCCCTGTCTCTACTATAAGATGTAATAA 1

RESULT 9

US-08-485-862B-65/c
Sequence 65, Application US/08485862B
Patent No. 5989838
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,862B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-485-862B-65

Query Match 0.3%; Score 76.2; DB 4; Length 105;
Best Local Similarity 82.9%; Pred. No. 7.3e-09;
Matches 87; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 9319 ATCCACGACTTTGGGAGCGGCGGCGGATCAGGATCAGGAGTGAGGACCATC 9378
Db 105 ATCCACGACTTTGGGAGCGGCGGCGGATCAGGATCAGGAGTGAGGACCATC 46
QY 9379 CTGCTTAAACAGATGAACCCCGTCTCTACTATAAAATACAAATA 9423
Db 45 CTGCCCAATATGGTGAACCCCTGTCTCTACTATAAGATGTAATAA 1

RESULT 10

US-08-787-739-65/c
Sequence 65, Application US/08787739
Patent No. 6027887
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 96
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 369 Pine Street, Suite 610
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/787,739
FILING DATE: 24-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/486,756
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/481,658
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,863
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,077
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.4
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-787-739-65


```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-450-673C-91

Query Match 0.2%; Score 67.6; DB 4; Length 84;
Best Local Similarity 89.0%; Pred. No. 6.5e-07;
Matches 73; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 9309 CACGCTGTAAATCCAGCACTTTGGAGCGCGGCGGATCAGAGGTTCAGGAGAT 9368
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 83 CACGCTGTAAATCCAGCACTTTGGAGCGCGGCGGATCAGAGGTTCAGGAGTT 24

QY 9369 GGAGACCATCCTGCTTAACACG 9390
      || |||| |||| |||| ||
Db 23 CGACACGAGCCTGATGAACATG 2

RESULT 14
US-08-450-673C-91
; Sequence 91, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-454-557C-69

Query Match 0.2%; Score 62.8; DB 3; Length 76;
Best Local Similarity 97.0%; Pred. No. 8.1e-06;
Matches 64; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10842 ACCATGCCCGCTAAATTTTGTATTTTAGTAGACAGAGGTTTCACCGTGTGGCCAGG 10901
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ACCACGCCCGCTAAATTTTGTATTTTAGTAGACAGAGGTTTCACCGTGTGGCCAGG 60

QY 10902 ATGTTTC 10907
      |||| ||
Db 61 ATGCTC 66
```

```
;
; TOPOLOGY: both
;
PCT-US95-17111A-91

Query Match 0.2%; Score 67.6; DB 6; Length 84;
Best Local Similarity 89.0%; Pred. No. 6.5e-07;
Matches 73; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 9309 CACGCTGTAAATCCAGCACTTTGGAGCGCGGCGGATCAGAGGTTCAGGAGAT 9368
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 83 CACGCTGTAAATCCAGCACTTTGGAGCGCGGCGGATCAGAGGTTCAGGAGTT 24

QY 9369 GGAGACCATCCTGCTTAACACG 9390
      || |||| |||| |||| ||
Db 23 CGACACGAGCCTGATGAACATG 2

RESULT 15
US-08-454-557C-69
; Sequence 69, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-454-557C-69

Query Match 0.2%; Score 62.8; DB 3; Length 76;
Best Local Similarity 97.0%; Pred. No. 8.1e-06;
Matches 64; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10842 ACCATGCCCGCTAAATTTTGTATTTTAGTAGACAGAGGTTTCACCGTGTGGCCAGG 10901
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ACCACGCCCGCTAAATTTTGTATTTTAGTAGACAGAGGTTTCACCGTGTGGCCAGG 60

QY 10902 ATGTTTC 10907
      |||| ||
Db 61 ATGCTC 66
```

Search completed: June 19, 2000, 03:08:43
Job time: 396853 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2000, 03:02:32 ; Search time 17971 Seconds
(without alignments)
-1569.860 Million cell updates/sec

Title: US-08-852-495C-2_COPY_112000_141000
Perfect score: 29001
Sequence: 1 TGTACTCTGGCTACCTCTG.....CTATAAAGTGAACAACCC 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 882769 seqs, -486395729 residues
Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_bal.*
- 2: gb_ba2.*
- 3: gb_om.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_pl1.*
- 8: gb_pl2.*
- 9: gb_pr1.*
- 10: gb_pr2.*
- 11: gb_pr3.*
- 12: gb_ro.*
- 13: gb_sts.*
- 14: gb_sy.*
- 15: gb_un.*
- 16: gb_v1.*
- 17: em_fun.*
- 18: em_hum1.*
- 19: em_hum2.*
- 20: em_in.*
- 21: em_om.*
- 22: em_or.*
- 23: em_ov.*
- 24: em_pat.*
- 25: em_ph.*
- 26: em_pl.*
- 27: em_ro.*
- 28: em_sts.*
- 29: em_sy.*
- 30: em_un.*
- 31: em_v1.*
- 32: gb_htg1.*
- 33: gb_htg2.*
- 34: gb_in1.*
- 35: gb_in2.*
- 36: em_ba1.*
- 37: em_ba2.*
- 38: em_hum3.*
- 39: em_hum4.*
- 40: gb_pr4.*
- 41: gb_htg3.*
- 42: gb_htg4.*
- 43: gb_htg5.*
- 44: gb_htg6.*

- 45: gb_htg7.*
- 46: em_htg1.*
- 47: em_htg2.*
- 48: em_htg3.*
- 49: em_hum5.*
- 50: gb_pl3.*
- 51: gb_pr5.*
- 52: gb_btg8.*
- 53: gb_btg9.*
- 54: gb_btg10.*
- 55: gb_btg11.*
- 56: gb_btg12.*
- 57: gb_btg13.*
- 58: gb_btg14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	99	0.3	108	10	HSLDLRN2
2	90.2	0.3	107	9	HUMALCE162
3	89.4	0.3	108	10	HSLDLRD1
4	89.4	0.3	108	10	HSLDLRD2
5	86.8	0.3	108	10	HSLDLRN2
6	85.4	0.3	103	9	HUMALCE221
7	83.2	0.3	108	10	HSLDLI12
8	82.6	0.3	108	11	HSU67803
9	81.2	0.3	107	9	HUMALCE162
10	79.8	0.3	103	9	HUMALCE221
11	79.4	0.3	108	10	HSLDLRD1
12	79.4	0.3	108	10	HSLDLRD2
13	76	0.3	103	13	HS8IC8R
14	75.6	0.3	110	11	HSU67807
15	75.2	0.3	103	13	HS8IC8R
16	74.4	0.3	104	9	HUMALCE272
17	74.2	0.3	108	9	HUMDID03M5
18	74.4	0.3	108	11	HSU67808
19	73.6	0.3	97	9	HUMLDLRA2
20	73.4	0.3	110	11	HSU67807
21	71.4	0.2	108	11	HSU67804
22	70.8	0.2	99	13	HUMUT7692A
23	70.8	0.2	108	10	HSLDLI12
24	70	0.2	95	13	HUMUT8002B
25	70.2	0.2	108	13	G32614
26	69.8	0.2	101	10	S79560
27	69.6	0.2	107	11	HSU67806
28	69.4	0.2	108	13	G43535
29	69	0.2	90	9	HUMLDLRFL
30	69.2	0.2	91	13	HUMUT8164A
31	68.4	0.2	79	10	S73203
32	68.2	0.2	101	10	S79560
33	67.8	0.2	95	10	HSSHPK1B
34	68	0.2	108	13	G43535
35	67.2	0.2	80	9	HUMBRKFAE
36	66.6	0.2	91	13	HUMUT8164A
37	66.8	0.2	94	9	HUMHGAL
38	66.2	0.2	100	9	HUMGALNSA
39	66.2	0.2	100	9	HUMGALNSA
40	65.2	0.2	108	9	HUMDID03M5
41	64.4	0.2	90	10	HSU19407
42	64.6	0.2	100	10	HSLAS27
43	64.4	0.2	110	9	HUMALCE43
44	64	0.2	90	9	HUMLDLRFL
45	64	0.2	97	9	HUMLDLRA2

ALIGNMENTS

```

RESULT 1
HSLDLRN2      108 bp      DNA      PRI      20-MAY-1992
LOCUS          Human LDL-receptor gene intron 14 fragment (normal gene).
DEFINITION     X05250
ACCESSION      X05250.1 GI:34337
VERSION        Alu repetitive sequence; low density lipoprotein receptor.
KEYWORDS       human.
SOURCE         Homo sapiens
ORGANISM       Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
REFERENCE      Primates; Catarrhini; Homnidae; Homo.
AUTHORS        Horsthemke,B., Beisiegel,U., Dunning,A., HAVINGA,J.R.,
TITLE          Unequal crossing-over between two alu-repetitive DNA sequences in
               the low-density-lipoprotein-receptor gene. A possible mechanism for
               the defect in a patient with familial hypercholesterolaemia
JOURNAL        Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE        87161901
COMMENT        See X05252 for deletion junction
               Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES       Location/Qualifiers
               source      1..108
                       /organism="Homo sapiens"
                       /db_xref="taxon:9606"
               intron      1..108
               /note="intron XIV fragment"
BASE COUNT     28 a 23 c 39 g 18 t
ORIGIN

Query Match      0.3%; Score 99; DB 10; Length 108;
Best Local Similarity 95.3%; Pred. No. 1e-07; Indels 0; Gaps 0;
Matches 102; Conservative 0; Mismatches 5;

Qy 27523 ACAAAATACAGCGGGTGGTGGCAGTGTGCTGTATCCAGCTACTCAGGAGGCTGAG 27582
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 ACAAAATAGCAGCGGTGGTGGCAGTGTGCTGTATCCAGCTACTCAGGAGGCTGAG 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 27583 GCAGAGAAATGCTTGACCCAGGAGCGGGAGGTTGGCAGTGAGCCGA 27629
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 GCAGGAGAAATGCTTGACCCAGGAGCGAGAGTTGGCAGTGAGCCGA 107
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 2
HUMALCEL62/c  107 bp ss-RNA  PRI  15-APR-1994
LOCUS          Human carcinoma cell-derived Alu RNA transcript, clone CE162.
DEFINITION     M87924
ACCESSION      M87924.1 GI:174871
VERSION        Alu repeat.
KEYWORDS       Homo sapiens male embryo carcinoma cDNA to other RNA.
SOURCE         Homo sapiens
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
REFERENCE      Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS        Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE          Alu RNA transcripts in human embryonal carcinoma cells. Model of
               post-transcriptional selection of master sequences
JOURNAL        J. Mol. Biol. (1992) In press
MEDLINE        112000
COMMENT        Location/Qualifiers
               source      1..107
                       /organism="Homo sapiens"
                       /db_xref="taxon:9606"
                       /cell_line="Wraera2D1"
                       /dev_stage="embryo"
                       /sex="male"
                       /tissue_type="carcinoma"
BASE COUNT     28 a 30 c 35 g 14 t
ORIGIN

```

```

Query Match      0.3%; Score 90.2; DB 9; Length 107;
Best Local Similarity 92.2%; Pred. No. 3.3e-06;
Matches 95; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 22252 TTTTCTGAGACGAGTCTTCTGTCTGCGCCAGGCTGGAATGAGTGGCACATCTCGGC 22311
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 107 TTTTCTGAGACGAGTCTTCTGTCTGCGCCAGGCTGGAATGAGTGGCACATCTCGGC 48
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 22312 TCAGTCAACCTCGCCTCCGGATTTCACGCCATTCTCTCCCTGCC 22354
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 47 TCAGTCAACCTCGCCTCCGGATTTCACGCCATTCTCTCTGCC 5

RESULT 3
HSLDLRD1/c    108 bp      DNA      PRI      20-MAY-1992
LOCUS          Human LDL-receptor mutated gene with intron 12 deletion junction.
DEFINITION     X05249
ACCESSION      X05249.1 GI:34335
VERSION        Alu repetitive sequence; low density lipoprotein receptor.
KEYWORDS       human.
SOURCE         Homo sapiens
ORGANISM       Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
REFERENCE      Primates; Catarrhini; Homnidae; Homo.
AUTHORS        Williamson,R., Beisiegel,U., Dunning,A., HAVINGA,J.R.,
               Horsthemke,B., and Humphries,S.
TITLE          Unequal crossing-over between two alu-repetitive DNA sequences in
               the low-density-lipoprotein-receptor gene. A possible mechanism for
               the defect in a patient with familial hypercholesterolaemia
JOURNAL        Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE        87161901
COMMENT        *source: hypercholesterol aemia
               See X05248 for corresponding normal gene sequence
               In the defective LDL-receptor gene the deletion occurred between two
               alu-repetitive sequences, that are in the same direction, the
               deletion eliminates exons 13 and 14 and changes the reading frame
               of the resulting spliced mRNA.
               Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES       Location/Qualifiers
               source      1..108
                       /organism="Homo sapiens"
                       /db_xref="taxon:9606"
                       /cell_type="blood leukocytes from a patient with familial"
               misc_feature 1..108
                       /note="deletion junction region intron 12/ intron 15"
BASE COUNT     20 a 40 c 20 g 28 t
ORIGIN

Query Match      0.3%; Score 89.4; DB 10; Length 108;
Best Local Similarity 89.7%; Pred. No. 4.6e-06;
Matches 96; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 27523 ACAAAATACAGCGGTGGTGGCAGTGTGCTGTATCCAGCTACTCAGGAGGCTGAG 27582
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 108 ACAAAATTAGCAGCGGTGGTGGCAGTGTGCTGTATCCAGCTACTCAGGAGGCTGAG 49
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 27583 GCAGAGAAATGCTTGACCCAGGAGCGGGAGGTTGGCAGTGAGCCGA 27629
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 48 GCAGGAAATGTTTGAACCCAGGAGCGAGAGTTGGTGAGGCCGA 2

RESULT 4
HSLDLRD2      108 bp      DNA      PRI      20-MAY-1992
LOCUS          Human LDL-receptor mutated gene with intron 14 deletion junction.
DEFINITION     X05251
ACCESSION      X05251.1 GI:34336
VERSION        Alu repetitive sequence; low density lipoprotein receptor.
KEYWORDS       human.
SOURCE         Homo sapiens
ORGANISM       Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

```


COMMENT	see X05249 for deletion function
DATA	data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES	Location/Qualifiers
source	1. .108
misc_feature	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	complement(<1. .65)
	/note="Alu repeat"
intron	1. .108
	/note="intron XII fragment"
BASE COUNT	21 a 38 c 20 g 29 t
ORIGIN	
Query Match	0.3%; Score 83.2; DB 10; Length 108;
Best Local Similarity	87.5%; Pred. No. 5.4e-05;
Matches	91; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Qy 27526	AAATCATGCCAGCGGTGGTGGCATGTGGCTCTTAATCCAGCTACTTCAGGAGGCTGAGGCA 27585
Db 105	AAATCATGCCGGCGTGGTGGCATGTGGCTCTTAATCCAGCTACTTAAGGAGGCTGAGGCA 46
Qy 27586	AGCAAAATGCTTGAACCCAGGAGCGGAGGTTCGAGTCAGGCCGA 27629
Db 45	GGAAATGGTTTGAACCCAGGAGCGAGGTGTGGTGAGGCCGA 2
RESULT 8	
HSU67803/c	
LOCUS	HSU67803 108 bp RNA PRI 01-AUG-1997
DEFINITION	Human small cytoplasmic Alu transcript.
ACCESSION	U67803
VERSION	U67803.1 GI:2289917
KEYWORDS	Alu.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS	Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 108)
JOURNAL	Saikh, T.H., Roy, A.M., Kim, J., Batzer, M.A. and Deininger, P.L.
MEDLINE	cDNAs derived from primary and small cytoplasmic Alu (scAlu)
REFERENCE	transcripts
AUTHORS	J. Mol. Biol. 271 (2), 222-234 (1997)
TITLE	2 (bases 1 to 108)
JOURNAL	Saikh, T.H., Kim, J., Batzer, M.A. and Deininger, P.L.
FEATURES	Direct Submission
source	Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
	Children's Hospital of Philadelphia, 1004F Abramson Research
	Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
repeat_region	Location/Qualifiers
	1. .108
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="TscAlu2"
	1. .108
	/note="scAlu"
	/rpt_family="Alu"
	/rpt_type="dispersed
BASE COUNT	23 a 39 c 30 g 16 t
ORIGIN	
Query Match	0.3%; Score 82.6; DB 11; Length 108;
Best Local Similarity	90.7%; Pred. No. 6.9e-05;
Matches	88; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 22422	GTAGAGACAGGTTTCACCGTGTAGCCGGATGGTCTCGATCTCCTGACCTCATGATCT 22481
Db 97	GTAGAGACGGGTTTCACCTGTAGCCAGATGGTCTCGATCTCCTGACCTCGTATCC 38
Qy 22482	GCCACCTTCACGCTCCCAAGTGGCTAGGATCACAGGC 22518

Best Local Similarity 87.9%; Pred. No. 0.00021;
Matches 87; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 22286 CTGGATCAGTGGACAAATCTCGGCTCAGTCAACCTCCGCTCCGGATTCAAGCCAT 22345

Db 103 CTGGAGTGAATGGACGACATCTCGGCTCAGTCAACCTCCGCTCCGGATTCAAGCCAT 44

Qy 22346 TCTCTGCTCTCAACCTCCGAGTAGCTGGGACCAAGGC 22384

Db 43 TCTCTGCTTAGCTTCCGTTAGCTGGGATTACAGC 5

RESULT 11

LOCUS HSLDLRD1 108 bp DNA PRI 20-MAY-1992
DEFINITION Human LDL-receptor mutated gene with intron 12 deletion junction.
ACCESSION X05249

VERSION X05249.1 GI:34335

KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Horsthemke, B., Beisiegel, U., Dunning, A., Havinga, J.R.,
Williamson, R. and Humphries, S.

TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia

JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)

MEDLINE 87161901

COMMENT *source: hypercholesterol aemia

See X05248 for corresponding normal gene sequence
In the defective LDL-receptor gene the deletion occurred between two
alu-repetitive sequences, that are in the same direction, the
deletion eliminates exons 13 and 14 and changes the reading frame
of the resulting spliced mRNA.

Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.

FEATURES

source Location/Qualifiers

1..108

/organism="Homo sapiens"

/db_xref="taxon:9606"

/cell_type="blood leukocytes from a patient with familial"

1..108

/note="deletion junction region intron 12/ intron 15"

20 a 40 c 20 g 28 t

BASE COUNT

ORIGIN

Query Match 0.3%; Score 79.4; DB 10; Length 108;

Best Local Similarity 84.8%; Pred. No. 0.00025;

Matches 89; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 22307 TCGGCTCACTGCAACCTCCGCTCCGGATTCAAGCCATTCCTCGCTCAACCTCCCGA 22366

Db 2 TCGGCTCACTGCAACCTCCGCTCCGGATTCAAGCCATTCCTCGCTCAACCTCCCGA 61

Qy 22367 GTAGTGGGACCAAGCGCCGCCACCAAGCCAGCTAAATTTT 22411

Db 62 GTAGTGGGATTACAGGACCTGCCACCAAGCCCTGGCTAAATTTT 106

RESULT 12

LOCUS HSLDLRD2/c 108 bp DNA PRI 20-MAY-1992
DEFINITION Human LDL-receptor mutated gene with intron 14 deletion junction.
ACCESSION X05251

VERSION X05251.1 GI:34336

KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 108)
Horsthemke, B., Beisiegel, U., Dunning, A., Havinga, J.R.,
Williamson, R. and Humphries, S.

TITLE

Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia

JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)

MEDLINE 87161901

COMMENT *source: hypercholesterol aemia

See X05250 for corresponding normal gene sequence
In the defective LDL-receptor gene the deletion occurred between two
alu-repetitive sequences, that are in the same direction, the
deletion eliminates exons 13 and 14 and changes the reading frame
of the resulting spliced mRNA.

Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.

FEATURES

source Location/Qualifiers

1..108

/organism="Homo sapiens"

/db_xref="taxon:9606"

/cell_type="blood leukocytes from a patient with familial"

1..108

/note="intron XIV fragment"

28 a 20 c 40 g 20 t

BASE COUNT

ORIGIN

Query Match 0.3%; Score 79.4; DB 10; Length 108;

Best Local Similarity 84.8%; Pred. No. 0.00025;

Matches 89; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 22307 TCGGCTCACTGCAACCTCCGCTCCGGATTCAAGCCATTCCTCGCTCAACCTCCCGA 22366

Db 107 TCGGCTCACTGCAACCTCCGCTCCGGATTCAAGCCATTCCTCGCTCAACCTCCCGA 48

Qy 22367 GTAGTGGGACCAAGCGCCGCCACCAAGCCAGCTAAATTTT 22411

Db 47 GTAGTGGGATTACAGGACCTGCCACCAAGCCCTGGCTAAATTTT 3

RESULT 13

LOCUS HS8IC8R/c

DEFINITION Human sequence tagged site 8IC8R DNA from 19q13.

ACCESSION X57789

VERSION X57789.1 GI:23938

KEYWORDS STS; myotonic dystrophy.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 103)

Aldridge, F.L.

Direct Submission

Submitted (12-FEB-1991) F.L. Aldridge, ICI Pharmaceuticals,

Alderley Park, Macclesfield, Cheshire, SK10 4TG, UK

REFERENCE 2 (bases 1 to 103)

AUTHORS Butler, R., Riley, J.H., Ogilvie, D.J., Anand, R., Buxton, J.,

Davies, J., Johnson, K. and Markham, A.F.

Two sequence-tagged sites defining the ends of a 380 kb YAC clone

from 19q13

JOURNAL Nucleic Acids Res. 19 (17), 4787 (1991)

MEDLINE 91367697

COMMENT See also X57788 for STS 8IC8L.

FEATURES

source Location/Qualifiers

1..103

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="19q13"

/germline

/clone_lib="YAC library: ICI"

/clone="8IC8"

BASE COUNT

ORIGIN

23 a 28 c 23 g 22 t 1 others

Db 48 AGACCATCTGGCCAAACAYAGGAAAACCTCATCTCTACAAAAAGACA 1

```
RESULT 2
X12095
ID X12095 standard; DNA; 108 BP.
AC X12095;
DE Human biallelic polymorphic DNA fragment TIGR-A003M18a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medication;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; US-030455.
PR 06-NOV-1996; US-030455.
PS (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA Hudson T, Lander ES, Wang D;
WPI; 98-286974/25.
DR New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1: Page 219; 31Opp: English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;

Query Match 0.2%; Score 68.6; DB 1; Length 108;
Best Local Similarity 82.8%; Pred. No. 0.013;
Matches 77; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

QY 12486 TGTATTTTGTAGAGAGCGGGTTTACCAATGTTGGCAGGCTGTCTCAAACTCCTGAC 12545
Db 1 TGTCTTTTGTAGAGATGAGGTTTCTCTGTGTTGGCAGGATGTTCTCGAATCCTGAC 60

QY 12546 CTCAGGTGATCCACTGCCTCGGCTCCCAAAA 12578
Db 61 TTCAAGTGATCCGTCTGCTTGGCTCCCAAAA 93

RESULT 3
T25009
ID T25009 standard; cDNA to mRNA; 108 BP.
AC T25009;
DE Human gene signature HUMGS07131.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PS (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.

Query Match 0.2%; Score 68.6; DB 1; Length 108;
Best Local Similarity 82.8%; Pred. No. 0.013;
Matches 77; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

QY 12486 TGTATTTTGTAGAGAGCGGGTTTACCAATGTTGGCAGGCTGTCTCAAACTCCTGAC 12545
Db 1 TGTCTTTTGTAGAGATGAGGTTTCTCTGTGTTGGCAGGATGTTCTCGAATCCTGAC 60

QY 12546 CTCAGGTGATCCACTGCCTCGGCTCCCAAAA 12578
Db 61 TTCAAGTGATCCGTCTGCTTGGCTCCCAAAA 93
```

```
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 1748; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 34 A; 31 C; 26 G; 15 T;

Query Match 0.2%; Score 63.4; DB 1; Length 108;
Best Local Similarity 73.8%; Pred. No. 0.082;
Matches 79; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 27591 ATTGCTTTGAACCCAGGAGCGGAGTTGTCAGTGCAGCGAATCGCCACTGCTCCAG 27650
Db 2 ATGCGCTGAGCCATGAGCGCAAGGCTGTCAGTGCCATGTCACGCCACTGNATTCCAG 61

QY 27651 CTTGGGTAAACAGACAGGAGCTCTGTTTCAAAAATAAATAATACATA 27697
Db 62 CCTGAGTGACAGACAGCAAGCCCTGTTGAAAACACACACACACANCAA 108

RESULT 4
T24892
ID T24892 standard; cDNA to mRNA; 100 BP.
AC T24892;
DE Human gene signature HUMGS06998.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PS (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 1720; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
```


PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
tissues

PS Claim 1: Page 1748; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-R26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 34 A; 31 C; 26 G; 15 T;

Query Match 0.2%; Score 58.6; DB 1; Length 108;
Best Local Similarity 71.0%; Pred. No. 0.45;
Matches 76; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

OY 22239 TTTTCTTTCTTCTGAGACGAGCTGTGCTCTGTCGCCAGCTGGAATCACTG 22298

DB 108 TTGTGTTGTTGTTTCAACAGGGTGTGCTCTGCTCACTAGGCTGGAATCACTG 49

OY 22299 GCACAATCTCGCTCACTGCACACTCCGCTCCCGGATTCACGCCAT 22345

DB 48 GCGTGACCATGGCTCACTGCAGCGCTTGCGCTCATGGCTCAGCGCAT 2

RESULT 8

ID T21566 standard; cDNA to mRNA; 87 BP.
AC T21566;
DE 03-AUG-1996 (first entry)
DE Human gene signature HUMGS02944.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
tissues

PS Claim 1: Page 914; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-R26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

SQ Sequence 87 BP; 35 A; 21 C; 16 G; 13 T;

Query Match 0.2%; Score 58; DB 1; Length 87;
Best Local Similarity 79.8%; Pred. No. 0.55;
Matches 67; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

OY 23100 GATCACTTGAGTCCAGGAGTTTGAGACCAGCTGTGCTCAACATGGGAAACCTCATCTCTA 23159
DB 1 GATCGCTTGANCCAGGAGTTTAAACACGCGCCGAGGAACATGGCGAAACCCCTCTTTA 60

OY 23160 CAAAAATAAAAAAATTTGTCAGG 23183

DB 61 CAAAAATACAGAAATNAGCCAAG 84

RESULT 9

T21566/c
ID T21566 standard; cDNA to mRNA; 87 BP.
AC T21566;
DE 03-AUG-1996 (first entry)
DE Human gene signature HUMGS02944.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
tissues

PS Claim 1: Page 914; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-R26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 87 BP; 35 A; 21 C; 16 G; 13 T;

Query Match 0.2%; Score 58; DB 1; Length 87;
Best Local Similarity 79.8%; Pred. No. 0.55;
Matches 67; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

OY 9689 CTTGGCTAATTTTCTATTTTAGGAGAGATGGGGTTTCCACCATGTTGGCCAGACTGGTC 9748

DB 84 CTTGGCTAATTTTCTATTTTAGGAGATGGGGTTTCCACCATGTTCTCTCGGCTGGTT 25

OY 9749 TCAAACTCCTGCGCTCAAGTGATC 9772

DB 24 TTAACCTCCTGGGNTCAAGCGATC 1

RESULT 10

T26828
ID T26828 standard; cDNA to mRNA; 108 BP.

T26828;
14-NOV-1996 (first entry)
Human gene signature HUMGS09078.
DE Human signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; Cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATSU) MATSUBARA K.
PI (OKUBO) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 2182; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 18 A; 33 C; 23 G; 28 T;

Query Match 0.2%; Score 58.2; DB 1; Length 108;
Best Local Similarity 77.5%; Pred. No. 0.52;
Matches 69; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 22461 GATCTCTGACCTCATGATCTGCCACCTCAGCCCTCCCAAGTGTAGGATCAGCGCAT 22520
|||||
1 GATCTCTGACCTCATGATCTGCCCGGTNTGCGCTCCCATAGTGTGGGTTACAGGCAT 60

QY 22521 GAGCCACCGCGCCGCTACTGACTTTT 22549
|||||
61 GAGCCACCGCGCGCTGTTTATTCT 89

RESULT 11
X12087/c
ID X12087 standard; DNA; 100 BP.
AC X12087;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment EST98276a.
KW polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease

CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-TUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match 0.2%; Score 57; DB 1; Length 100;
Best Local Similarity 72.7%; Pred. No. 0.8;
Matches 72; Conservative 1; Mismatches 26; Indels 0; Gaps 0;

QY 27408 GTGGCTCACCTGTAAATCCAGCACTTTGGGAGCCCAAGGTAAGCAGATCACTTGAGT 27467
|||||
99 GTGGCTCACCTGTAAATCCCTGGCACTTTAGGAGGCTTAGGAAGGAGGATCTTTGAAC 40

QY 27468 CAGGAGTGTAGAGCAGCACTCTGGCCCAACATAGTGAAATC 27506
|||||
39 CAGGAGCTCAGACCAKCTGGGAACATAGCAAGCATC 1

RESULT 12
X12085/c
ID X12085 standard; DNA; 100 BP.
AC X12085;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment EST98276c.
KW polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease

PS Claim 1; Page 218; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-TUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular

```
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases. 25 C; 22 G; 30 T;
SQ Sequence 100 BP; 22 A; 25 C; 22 G; 30 T;

Query Match 0.2%; Score 57; DB 1; Length 100;
Best Local Similarity 72.7%; Pred. No. 0.8;
Matches 72; Conservative 1; Mismatches 26; Indels 0; Gaps 0;

Qy 27408 GTGGCTCACACTGTAATCCAGACACTTTGGGAGGCCAAGTAAGCAGATCACTTGAGCT 27467
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 99 GTGACTCACACTGTAATCCAGACACTTTGGGAGGCCAAGTAAGCAGATCACTTGAGCT 27467
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 27468 CAGGAGTTAGACAGACGCTGCGCAACATAGTGAACATC 27506
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 39 CAGGAGCTAAGACACCTCTGGGAACATAGCAGACATC 1

RESULT 13
T20977
ID T20977 standard; cDNA to mRNA; 103 BP.
AC T20977;
DE 24-JUL-1996 (first entry)
DT Human gene signature HUMGS02180.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATSU) MATSUBARA K.
PA (OKUBU) OKUBO K.
PI Matsubara K, Okubo K;
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 758-759; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-R26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 103 BP; 22 A; 27 C; 21 G; 31 T;

Query Match 0.2%; Score 57.2; DB 1; Length 103;
Best Local Similarity 77.3%; Pred. No. 0.75;
Matches 68; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 12430 TCCCTCCCAAGTAGCTGGGACTACAGGTGATACCAACCGGCCCTGCTAATTTTGTGA 12489
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13 TCCACCTCCCAAGTAGCTGGGCTAGGTGTGCCACCACTGTCACGCTGATTTTNGTA 72
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 12490 TTTTGTATAGACGGGGTTTCCACCATG 12517
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 73 TTTTNGTAGGACAGATATTCTCCCATG 100
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 14
X12087
ID X12087 standard; DNA; 100 BP.
AC X12087;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment EST98276a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC spherocytosis, von Willebrand's disease, tuberculous scleritis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases. 25 C; 22 G; 31 T;
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match 0.2%; Score 56.4; DB 1; Length 100;
Best Local Similarity 75.0%; Pred. No. 0.99;
Matches 69; Conservative 1; Mismatches 22; Indels 0; Gaps 0;

Qy 3246 GCCATGTTGGCCAGCGCTGATCCGAACTCCTGATTTCTGGTAAATCCGCCGCTCAGCCT 3305
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8 GCTATGTTTCCAGCGTGTCTTTGAGCTCCTGGTTTCAACAACATCTCTCTTCTAAGCCT 67
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 3306 CTTAAAGTGCTTGAATTACAGCGGTGAGTCAC 3337
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 68 CCTAAAGTGCCAGGATTATAGGTGTGAGTCAC 99
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
X12086
ID X12086 standard; DNA; 100 BP.
AC X12086;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment EST98276b.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
```

PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tubercous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match 0.2%; Score 56.4; DB 1; Length 100;
Best Local Similarity 75.0%; Pred. No. 0.99;
Matches 69; Conservative 1; Mismatches 22; Indels 0; Gaps 0;
QY 3246 GCCATGTTGGCCAGGTGATCCGGAACCTCGTATTTCTGGTAATCGCCCGCTCAGCCT 3305
Db 8 GCTATGTTTCCAGGATGCTTTGAGCTCCTGGTTCAAAACAACTCTCTTCTTACGCT 67
QY 3306 CTTAAAGTCTTGAATTACAGCGGTGAGTCAC 3337
Db 68 CCTAAAGTCCAGGATTATAGGTGTGAGTCAC 99

Search completed: June 19, 2000, 11:47:50
Job time: 427781 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: June 18, 2000, 22:02:37 ; Search time 8514.75 Seconds
(without alignments)
13805.165 Million cell updates/sec

Title: US-08-852-495C-2_COPY_112000_141000
Perfect score: 29001
Sequence: 1 TGTACTCTGGCTACCTCTG.....CTATAACTGGAACAACCC 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues
Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

- EST:*
- 1: em_est1:*
 - 2: em_est2:*
 - 3: em_est3:*
 - 4: em_est4:*
 - 5: em_est5:*
 - 6: em_est6:*
 - 7: em_est7:*
 - 8: em_est8:*
 - 9: em_est9:*
 - 10: em_est10:*
 - 11: em_est11:*
 - 12: em_est12:*
 - 13: em_est13:*
 - 14: em_est14:*
 - 15: em_est15:*
 - 16: em_est16:*
 - 17: em_est17:*
 - 18: em_est18:*
 - 19: em_est19:*
 - 20: gb_est1:*
 - 21: gb_est2:*
 - 22: gb_est3:*
 - 23: gb_est4:*
 - 24: gb_est5:*
 - 25: gb_est6:*
 - 26: gb_est7:*
 - 27: gb_est8:*
 - 28: gb_est9:*
 - 29: gb_est10:*
 - 30: gb_est11:*
 - 31: gb_est12:*
 - 32: gb_est13:*
 - 33: gb_est14:*
 - 34: gb_est15:*
 - 35: gb_est16:*
 - 36: gb_est17:*
 - 37: gb_est18:*
 - 38: gb_est19:*
 - 39: gb_est20:*
 - 40: gb_est21:*
 - 41: gb_est22:*
 - 42: gb_est23:*
 - 43: gb_est24:*
 - 44: gb_est25:*

- 45: gb_est26:*
- 46: gb_est27:*
- 47: gb_est28:*
- 48: gb_est29:*
- 49: gb_est30:*
- 50: gb_est31:*
- 51: gb_est32:*
- 52: em_est20:*
- 53: em_est21:*
- 54: em_est22:*
- 55: em_est23:*
- 56: em_est24:*
- 57: em_est25:*
- 58: em_est26:*
- 59: gb_est33:*
- 60: gb_est34:*
- 61: gb_est35:*
- 62: gb_est36:*
- 63: gb_est37:*
- 64: gb_est38:*
- 65: em_est27:*
- 66: em_est28:*
- 67: em_est29:*
- 68: em_est30:*
- 69: gb_est39:*
- 70: gb_est40:*
- 71: gb_est41:*
- 72: gb_est42:*
- 73: gb_est43:*
- 74: gb_est44:*
- 75: em_est31:*
- 76: em_est32:*
- 77: em_est33:*
- 78: em_est34:*
- 79: gb_est45:*
- 80: gb_est46:*
- 81: gb_est47:*
- 82: gb_gss1:*
- 83: gb_gss2:*
- 84: gb_gss3:*
- 85: gb_gss4:*
- 86: em_gss1:*
- 87: em_gss2:*
- 88: em_gss3:*
- 89: em_gss4:*
- 90: gb_gss5:*
- 91: gb_gss6:*
- 92: gb_gss7:*
- 93: gb_gss8:*
- 94: gb_gss9:*
- 95: em_gss5:*
- 96: em_gss6:*
- 97: em_gss7:*
- 98: em_gss8:*
- 99: em_gss9:*
- 100: em_gss10:*
- 101: em_gss11:*
- 102: gb_gss10:*
- 103: gb_gss11:*
- 104: em_gss12:*
- 105: gb_gss12:*
- 106: gb_gss13:*
- 107: gb_gss14:*
- 108: gb_gss15:*
- 109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result % Query

No.	Score	Match	Length	DB	ID	Description
1	93.2	0.3	106	37	AA703692	ag81a10.r
2	91.4	0.3	109	30	AA243009	zr25h02.s
3	90.4	0.3	106	30	AA250812	z806a05.s
c 4	88.6	0.3	103	84	B48914	RPC111-4A12
5	88.2	0.3	101	39	AA835205	ak64h01.s
c 6	88.2	0.3	110	30	AA244245	nc07a04.s
c 7	87.8	0.3	101	33	AA381369	EST94442
c 8	87.8	0.3	108	34	B65160	CIT-HSP-201
c 9	86.6	0.3	102	36	AA654562	nt75f10.s
c 10	86.2	0.3	107	35	AA565533	nk42b11.s
c 11	86.4	0.3	107	39	AA828124	od71a07.s
c 12	86.6	0.3	110	94	AQ003188	RPC111-1D
13	86	0.3	102	36	AA654562	nt75f10.s
c 14	86	0.3	103	108	AQ582186	RPC1-11-4
c 15	85.4	0.3	103	108	AQ535244	RPC1-11-3
c 16	85.2	0.3	109	94	AQ028426	CIT-HSP-2
c 17	85	0.3	110	30	AA244245	nc07a04.s
c 18	85.2	0.3	110	39	AA897366	am06h02.s
19	84.4	0.3	105	21	T94466	ve35b02.r1
c 20	84.6	0.3	107	103	AQ240182	CIT-HSP-2
c 21	84.2	0.3	106	108	AQ344957	CITBI-EI-
c 22	84.4	0.3	110	106	AQ386882	RPC111-13
c 23	83.8	0.3	103	94	AQ028649	CIT-HSP-2
c 24	83.8	0.3	103	108	AQ535244	RPC1-11-3
25	83.6	0.3	106	63	AI991750	wt48e01.x
c 26	83.6	0.3	106	63	AI991750	wt48e01.x
c 27	83.6	0.3	109	84	B17434	345K2.TVB C
c 28	83.2	0.3	96	92	AQ936334	RPC1-11-S
c 29	83.2	0.3	104	105	AQ321855	RPC111-11
c 30	83.4	0.3	109	84	B17434	345K2.TVB C
31	83	0.3	101	33	AA381369	EST94442
c 32	82.8	0.3	103	38	AA807640	nx08b05.s
c 33	82.2	0.3	103	108	AQ584425	RPC1-11-4
c 34	82.2	0.3	105	109	AQ637292	RPC1-11-4
c 35	82.4	0.3	108	84	B32951	HS-1016-A1-
c 36	81.6	0.3	104	105	AQ321855	RPC111-11
c 37	81.8	0.3	106	30	AA250812	z806a05.s
c 38	82	0.3	106	38	AA812141	eb48h02.s
c 39	82	0.3	106	50	AI700000	tt36a10.x
c 40	81.6	0.3	105	30	AA218889	zq15d04.s
c 41	81	0.3	105	28	AA078003	7H12D08 C
c 42	81.2	0.3	110	106	AQ386882	RPC111-13
c 43	80.6	0.3	107	33	AA385808	EST9495
c 44	80.2	0.3	101	94	AQ076649	CIT-HSP-2
45	80.2	0.3	101	105	AQ260734	CITBI-EI-

ALIGNMENTS

```

RESULT 1
AA703692
LOCUS
DEFINITION ag81a10.r1 StrataGene hNT neuron (#937233) Homo sapiens cDNA clone
IMAGE:1140858 5' similar to contains Alu repetitive element; , mRNA
sequence.
ACCESSION AA703692
VERSION AA703692.1 GI:2713610
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 106)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le.N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1397630.

```

```

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 53.
Location/Qualifiers
1..106
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:1140858"
/clone_lib="Stratagene hNT neuron (#937233)"
/lab_stage="hNT neurons"
/dev_host="SOLR (kanamycin resistant)"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT.
Differentiated, post mitotic hNT neurons. Average insert
size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'
GAATCGGCACGAG 3' -3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT 19 a 29 c 29 g 29 t
ORIGIN
Query Match 0.3%; Score 93.2; DB 37; Length 106;
Best Local Similarity 92.5%; Pred. No. 0.15;
Matches 98; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 22416 TTTTATTAGACAGAGGTTTCCGCTGTTAGCCGGGATGCTCTCGATCTCCTGACCTCA 22475
|||||
Db 1 TTTTATTAGACAGAGGTTTCCGCTGTTAGCCGGGATGCTCTCGATCTCCTGACCTCG 60
|||||
QY 22476 TGATCTGCCACCCAGCAGCTCCCAAGTGTGATGATCAGAGCGCATG 22521
|||||
Db 61 TGATCTGCCCGCCTCAGCCTCCCAAGTGTGATGATCAGAGCGCATG 106
|||||
RESULT 2
AA243009
LOCUS
DEFINITION zr25h02.s1 StrataGene NT2 neuronal precursor 937230 Homo sapiens
cDNA clone IMAGE:664467 3' similar to contains Alu repetitive
element; contains element LTR1 repetitive element ; , mRNA sequence.
ACCESSION AA243009
VERSION AA243009.1 GI:1873869
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 109)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le.N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)
On Dec 3, 1996 this sequence version replaced gi:1126869.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1127 Std Error: 0.00
Seq primer: -44m13 fwd. ET from Amersham
High quality sequence stop: 102.

```

FEATURES
source

Location/Qualifiers
1. .109
/organism="Homo sapiens"
/db_xref="CDB:5426481"
/db_xref="taxon:9606"
/clone="IMAGE:664467"
/tissue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: brain; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: oligo dt. Uninduced, exponentially growing neuroepithelial cells (Ntera-2/cl.D1). Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3'"

BASE COUNT
ORIGIN

19 a 30 c 30 g 30 t

Query Match 0.3%; Score 91.4; DB 30; Length 109;
Best Local Similarity 89.9%; Pred. No. 0.24; Indels 0; Gaps 0;
Matches 98; Conservative 0; Mismatches 11;

Qy 22413 GTATTTTAGTAGACAGGGTTTCACCGTGTACCGGGATGGTCTCGATCTCCTGACC 22472
|||||
Db 1 GTATTTTAGTAGACAGGGTTTCACCGTGTACCGAGGATGGTCTGATCTCCTTACC 60

Qy 22473 TCATGATCTGCCACCTCAGCGCTCCCAAGTGTAGTAGTACAGGCATG 22521
|||||
Db 61 TCGTGATCGCGCCACCTCGCGCTCCCAAGTGTGGGATTACAGGCGTG 109

RESULT 3
AA250812

LOCUS AA250812 106 bp mRNA EST 15-AUG-1997
DEFINITION zs06a05.s1 NCI-CGAP GCB1 Homo sapiens cDNA clone IMAGE:684368 3' similar to contains Alu repetitive element; contains element MER22 repetitive element ;, mRNA sequence.

ACCESSION AA250812

VERSION AA250812.1 GI:1885774

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 106)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT On Sep 12, 1996 this sequence version replaced gi:1407356.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 537 Std Error: 0.00

Seq primer: -41m13 fwd. Et from Amersham

High quality sequence stop: 87.

Location/Qualifiers

1. .106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:684368"
/clone_lib="NCI-CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"

/note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman

FEATURES
source

(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGGGGGGGCTCATTTTTTTTTTTTTTTT-3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 20 a 28 c 31 g 27 t
ORIGIN

Query Match 0.3%; Score 90.4; DB 30; Length 106;
Best Local Similarity 94.0%; Pred. No. 0.31;
Matches 94; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 12497 TAGAGACGGGGTTTCACCATGTTGGCCAGGCTGCTCAAACTCTTGACCTCAGGTGATC 12556
|||||
Db 2 TAGAGACGGGGTTTCACCATGTTGGCCAGGCTGCTCAAACTCTTGACCTCAGGTGATC 61

Qy 12557 CACCTGCCTCGCCCTCCCAAAATGCTGAGATTACAGGTGT 12596
|||||

Db 62 CACTTGCCTTGCCCTCCCAAAAGTGTGGGATTACAGGTGT 101

RESULT 4

B48914/c

LOCUS B48914 103 bp DNA

DEFINITION RPC111-4A12.TP RPCI-11 Homo sapiens genomic clone RPCI-11-4A12, genomic survey sequence.

ACCESSION B48914

VERSION B48914.1 GI:2601151

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 103)
Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and Venter, J.C.

TITLE Use of BAC End Sequences for Sequence-Ready Map Building

JOURNAL Unpublished (1997)

COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers
1. .103
/organism="Homo sapiens"
/db_xref="GDB:7501163"
/db_xref="taxon:9606"
/clone="RPCI-11-4A12"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; RPC111 Human Male BAC Library"

FEATURES
source

30 a 28 c 30 g 15 t
ORIGIN

Query Match 0.3%; Score 88.6; DB 84; Length 103;
Best Local Similarity 91.3%; Pred. No. 0.5;
Matches 94; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY	22412	TGTAATTTT	TAGTAGACACAGGGTTT	CACCGTGTT	TAGCCGGGATGGTCTCGATCTCTGAC	22471
Db	103	TGTAATTTT	TAGTAGACACGGGGTTT	CACCGTTTT	TAGCCGGGATGGTCTCGATCTCTGAC	44
QY	22472	CTCATGATCTGCCACCT	CAGCCTCCCAAGT	GCTAGGATCAC	22514	
Db	43	CTGCTGATCCGCCCGCT	TCGGCGTCCCAAGT	GCTGGGGTTAC	1	

RESULT	5					
LOCUS	AA835205					23-FEB-1998
DEFINITION	ak64h01.s1 Barstead pancreas HPLRB1 Homo sapiens cDNA clone	101 bp	mRNA	EST		
	IMAGE:1412689 3' similar to contains Alu repetitive element; contains element KER repetitive element ; mRNA sequence.					

ACCESSION AA835205
VERSION AA835205.1 GI:2908933
KEYWORDS EST.
SOURCE human.

SOURCE	ORGANISM	Human?
	<i>Homo sapiens</i>	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	

REFERENCE	AUTHORS	TITLE
1	(bases 1 to 101)	
	Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Thesling, B., White, F., Wylie, T., Waterston, R. and Wilson, R.	WashU-NCI human EST project

JOURNAL
Unpublished (1997)
On Nov 29, 1993 this sequence version replaced q1:636191.
COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m13 fwd. ET from Amersham.

```

seq primer: 40m131NW.121 from Amersham.
Location/Qualifiers
1. .101
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1412689"
/clone_lib="Barstead pancreas HPLRB1"
/sex="female"
/dev_stage="adult, 34 years"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Site_1: EcoRI; Site_2: NotI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5',
TGTTACGAATCTCAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTTTT
3']"; double-stranded cDNA was ligated to Eco RI adaptors
[AAATCGATCCTTGG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT7T3 vector.
Library constructed by Bob Barstead. "
14 a 36 c 27 g 24 t
BASE COUNT
ORIGIN

```

Query Match 0.3%; Score 88.2; DB 39; Length 101;
Best Local Similarity 92.1%; Pred. No. 0.56;
Matches 93: Conservative 0; Mismatches 8; Indels 0;
Gaps 0;

QY 22257 TGAGACGGAGTCTTGTTGTTCTGTGCCCCAGGCTGGAATGCAGTGGCACAAATCTCGGCTCACT 22316
Db 1 TGAGACGGAGTCTCACTCTGTGCCCAGGCTGGAGTCAGTGGCTGATCTCGGCTCACT 60

Qy	22317	GCAACCTCGGCTCCGGATTACGGCAATCTCCTGCCTCA	22357
Db	61	GCAAGCTCGGCTCCGGGTTACGGCAATCTCCTGCCTCA	101

RESULT	6			
AA244245/c				
LOCUS	AA244245	110 bp	mRNA	EST
DEFINITION	nc07a04.s1 NCI_CGAP_Prl Homo sapiens cDNA clone IMAGE:1007406 similar to contains Alu repetitive element;; mRNA sequence.			20-AUG-1997

ACCESSION AA244245
VERSION AA244245.1
KEYWORDS EST.
SOURCE human.

SOURCE Homalini.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 110)

1 (bases 1 to 110)
 NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anat.
 Tumor Gene Index

JOURNAL
COMMENT

Unpublished (1997)
On Jan 24, 1995 this sequence version replaced g1:634306.
Contact: Robert Strausberg, Ph.D.

Tel.: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaquil,
M.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution Information Center
found through the I.N.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbbr/image/image.html

Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 90.

FEATURES	source
margin	quantitative
Location/Qualifiers	
1..1110	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone="IMAGE:1007406"	
/clone_lib="NCI_CGAP_Pri"	
/sex="Male"	
/dev_stage="45 years old"	
/lab_host="DH10B"	
/notes="PAMP10; SI"	
/vector="pGEMT"	
strand	cdna was primed with

strand cDNA was primed with oligo(dT)₁₇ on 20 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected, histologically normal prostate epithelial cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman.

BASE COUNT		ORIGIN	
17 a	26 c	28 g	38 t
considered of Saudi Arabian:			1 others

Query Match 0.3%; Score 88.2; DB 30; Length 110;
Best Local Similarity 87.3%; Pred. No. 0.54;
Matches 96; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 15415 TGAGGCAGGAGAATCGCTTTGAACCCAGCAGGCCAGAGGTTCAGTGGCCAAGTTCCTGTC 15474
 | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||
Dbb 110 TGAGGCAGGAGAATCTCTTGAAACCAGCAGGCCAGAGGTTGCAGTGAGCCAAGANTCTGCC 51

QY 15475 ACTGCACCCACCTGGGCGACAGCGAGACTTCGTCTCAAAAAACAA 15524
||||||| ||| ||||||| ||||| | ||||| | ||||||| |||
Db 50 ACTGCACCTCCAGCTGGGCAACAGATCAAGACTCCATCTCAAAAAAAA 1


```

RESULT 7
AA381369/c
LOCUS
DEFINITION
  EST94442 Activated T-cells 1 mRNA EST 21-APR-1997
  EST containing Alu repeat, mRNA sequence.
ACCESSION
AA381369
VERSION
AA381369.1 GI:2033689
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 101)
AUTHORS
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Pallanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Otterback,T.R., Weidman,J.F., Li,Y.,
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillon,P.J., Fannon,M.K., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
12140200
JOURNAL
MEDLINE
COMMENT
On Sep 12, 1996 this sequence version replaced gi:1407448.
Other ESTs: THCI70052
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlavet@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tldb/hgi/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
  1..101
  /organism="Homo sapiens"
  /db_xref="ATCC (inhost):185728"
  /db_xref="taxon:9606"
  /clone_lib="Activated T-cells 1"
  /cell_type="T-lymphocyte"
  /dev_stage="adult"
  /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT
18 a 36 c 20 g 25 t 2 others
ORIGIN
Query Match 0.3%; Score 87.8; DB 33; Length 101;
Best Local Similarity 91.1%; Pred. No. 0.62;
Matches 92; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 15362 AAAATTAGCTGGGTGGTGGCGGCACCTGTAATCCAGCTAATCAGGAGCTGAGGCA 15421
|||||
Db 101 AAAATTAGCTGGGAGTGGTGGCGGCGCTGTAATCCAGCTACTCAGGAGCTGAGGCA 42
|||||

QY 15422 GGAGATCGCTTGACCCAGGAGGAGGAGGTTGCAGTGAGC 15462
|||||
Db 41 GGANAATTGCTTGAACCCAGGAGGAGGAGGTTGCAATGAGC 1

RESULT 8
B65160
LOCUS
DEFINITION
  CIT-HSP-2017G2.TRB CIT-HSP Homo sapiens genomic clone 2017G2,
  genomic survey sequence.
ACCESSION
B65160
VERSION
B65160.1 GI:2639138
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 108)
AUTHORS
Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building
Unpublished (1997)
Other GSSs: CIT-HSP-2017G2.TFB
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/hungen/bac\_end\_search/bac\_end\_search.html
Seq primer: M13 Reverse
Class: BAC ends.
Location/Qualifiers
  1..108
  /organism="Homo sapiens"
  /db_xref="GDB:7043860"
  /db_xref="taxon:9606"
  /clone="2017G2"
  /clone_lib="CIT-HSP"
  /sex="Male"
  /cell_type="Sperm"
  /note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT
26 a 27 c 34 g 21 t
ORIGIN
Query Match 0.3%; Score 87.8; DB 84; Length 108;
Best Local Similarity 88.8%; Pred. No. 0.6;
Matches 95; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 27541 TGGTGGCATGTCCTGTATCCAGCTACTCAGGAGGCTGAGGCAAGAATCTTGAA 27600
|||||
Db 1 TGGTGGCATGTCCTGTATCCAGCTACTCAGGAGGCTGAGGCAAGAATCTTGAA 60
|||||

QY 27601 CCCAGGCGCGGTGTCAGTGAGCCGAATCGCCCACTGCACTC 27647
|||||
Db 61 CCCGGAGGTGGAGGTTCAGTGAGCCGAATCATACCACATGCACAC 107
|||||

RESULT 9
AA654562/c
LOCUS
DEFINITION
  nt75f10.s1 NCI-CCAP-Pr3 Homo sapiens cDNA clone IMAGE:1204363
  similar to contains Alu repetitive element;contains element MER22
  repetitive element ;, mRNA sequence.
ACCESSION
AA654562
VERSION
AA654562.1 GI:2590716
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens

```


5,000-10,000 microdissected cells histologically-determined to be fully malignant prostate cancer cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Kriman."									
BASE COUNT	22 a	32 c	27 g	21 t					
ORIGIN									
Query Match	0.38; Score 86; DB 36; Length 102;								
Best Local Similarity	90.2%; Pred. No. 0.97;								
Matches	92; Conservative	0; Mismatches	10; Indels	0; Gaps	0;				
QY 12500	AGACGGGGTTTCACCATCTTGGCCAGGCTGGTCTCTCAAACTCCTGACCTCAGGTGATCCAC 12559								
Db 1	AGACGAGGTTTCACCATCTTGGCCAGGCTGGTCTCTCAAACTCCGACCTCAGGTATCCGC 60								
QY 12560	CTGCTCTGGCCCTCCCAAAATGCTGAGATTACAGGTGTGAGCC 12601								
Db 61	CCACCTCGGCTCCCAAAAGTCTGGGATTACAGGAGTGAGCC 102								
RESULT 14									
AQ582186									
LOCUS	AQ582186	103 bp	DNA	GSS	07-JUN-1999				
DEFINITION	RPCI-11-451A15.TJ RPCI-11 Homo sapiens genomic clone								
	RPCI-11-451A15, genomic survey sequence.								
ACCESSION	AQ582186								
VERSION	AQ582186.1 GI:5009296								
KEYWORDS	GSS.								
SOURCE	human.								
ORGANISM	Homo sapiens								
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;								
	Eutheria; Primates; Catarrhini; Hominidae; Homo.								
REFERENCES	1 (bases 1 to 103)								
AUTHORS	Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and								
	Venter,J.C.								
TITLE	Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready								
	Map Building								
JOURNAL	Unpublished (1997)								
COMMENT	On Feb 19, 1999 this sequence version replaced gi:4146076.								
	Other GSSs: RPCI-11-451A15.TV								
	Contact: Shaying Zhao, William Nierman, Mark Adams								
	Department of Eukaryotic Genomics								
	The Institute for Genomic Research								
	9712 Medical Center Dr., Rockville, MD 20850								
	Tel: 301 838 0200								
	Fax: 301 838 0208								
	Email: hbe@tigr.org								
	Clones are derived from the human BAC library RPCI-11. For BAC								
	library availability, please contact Pieter de Jong								
	(pieter@edj.med.buffalo.edu). Clones may be purchased from								
	BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from								
	Research Genet cs (info@resgen.com). BAC end search page:								
	http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.								
	Seq primer: SP6								
Class:	BAC ends.								
FEATURES	Location/Qualifiers								
source	1..103								
	/organism="Homo sapiens"								
	/db_xref="GDB:7672814"								
	/db_xref="taxon:9606"								
	/clone="RPCI-11-451A15"								
	/clone_lib="RPCI-11"								
	/sex="Male"								
	/cell_type="Lymphocytes"								
	/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;								
	RPC111 Human Male BAC Library"								
BASE COUNT	19 a	36 c	25 g	22 t	1 others				

Db 43 TGACCTCAAGTGATCTGCCCGTCTTGGCCTCCCAAAGTCTGG 1

Search completed: June 19, 2000, 06:28:49
Job time: 409896 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2000, 03:08:43 ; Search time 372.76 seconds
(without alignments)
10112.932 Million cell updates/sec

Title: US-08-852-495C-2_COPY_112000_141000
Perfect score: 29001
Sequence: 1 TGTACTGTGCTACCTCTG.....CTATAACTGGAACAACCC 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCITUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78.4	0.3	105	4	US-08-481-658B-65
2	78.4	0.3	105	4	US-08-477-504A-65
3	78.4	0.3	105	4	US-08-486-756A-65
4	78.4	0.3	105	4	US-08-485-862B-65
5	78.4	0.3	105	5	US-08-787-739-65
6	68	0.2	105	4	US-08-481-658B-65
7	68	0.2	105	4	US-08-477-504A-65
8	68	0.2	105	4	US-08-486-756A-65
9	68	0.2	105	4	US-08-485-862B-65
10	68	0.2	105	5	US-08-787-739-65
11	60.4	0.2	78	3	US-08-454-557C-70
12	60.4	0.2	78	4	US-08-340-426D-70
13	60.4	0.2	78	4	US-08-450-673C-70
14	60.4	0.2	78	6	PCT-US95-17111A-70
15	58.8	0.2	78	3	US-08-454-557C-70
16	58.8	0.2	78	4	US-08-340-426D-70
17	58.8	0.2	78	4	US-08-450-673C-70
18	58.8	0.2	78	6	PCT-US95-17111A-70
19	59	0.2	84	3	US-08-454-557C-91
20	59	0.2	84	4	US-08-340-426D-91
21	59	0.2	84	4	US-08-450-673C-91
22	59	0.2	84	6	PCT-US95-17111A-91
23	53.2	0.2	76	3	US-08-454-557C-69
24	53.2	0.2	76	4	US-08-340-426D-69
25	53.2	0.2	76	4	US-08-450-673C-69
26	53.2	0.2	76	6	PCT-US95-17111A-69
27	53.4	0.2	85	3	US-08-454-557C-92

Sequence 92, Appl
Sequence 92, Appl
Sequence 92, Appl
Sequence 91, Appl
Sequence 91, Appl
Sequence 91, Appl
Sequence 91, Appl
Sequence 57, Appl
Sequence 57, Appl
Sequence 57, Appl
Sequence 57, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 60, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-65
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.3%; Score 78.4; DB 4; Length 105;
Best Local Similarity 84.6%; Pred. No. 2.4e-08;

Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 21613 TTTTGTATTTTCTAGTAAGACGGGTTTCCACCATGTTGGTCAGGCTGTCTCCAACCTCC 21672

Db 2 TTTTACATCTTTAGTAGAGACGGGTTTCCACCATGTTGGCCAGGCTGTCTCCAACCTCC 61

Qy 21673 TGACCTCATGATCTGCCACCTTGCCCTCCCAAAAGTGTGGGAT 21716

Db 62 TGACCTTGATCCACCAGCCTCGGCCCTCCCAAAAGTGTGGGAT 105

RESULT 2

US-08-477-504A-65

; Sequence 65, Application US/08477504A

; Patent No. 5972353

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,504A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3D

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-477-504A-65

Query Match 0.3%; Score 78.4; DB 4; Length 105;

Best Local Similarity 84.6%; Pred. No. 2.4e-08;

Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 21613 TTTTGTATTTTCTAGTAAGACGGGTTTCCACCATGTTGGTCAGGCTGTCTCCAACCTCC 21672

Db 2 TTTTACATCTTTAGTAGAGACGGGTTTCCACCATGTTGGCCAGGCTGTCTCCAACCTCC 61

Qy 21673 TGACCTCATGATCTGCCACCTTGCCCTCCCAAAAGTGTGGGAT 21716

Db 62 TGACCTTGATCCACCAGCCTCGGCCCTCCCAAAAGTGTGGGAT 105

RESULT 3

US-08-486-756A-65

; Sequence 65, Application US/08486756A

; Patent No. 5981711

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,756A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-486-756A-65

Query Match 0.3%; Score 78.4; DB 4; Length 105;

Best Local Similarity 84.6%; Pred. No. 2.4e-08;

Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 21613 TTTTGTATTTTCTAGTAAGACGGGTTTCCACCATGTTGGTCAGGCTGTCTCCAACCTCC 21672

Db 2 TTTTACATCTTTAGTAGAGACGGGTTTCCACCATGTTGGCCAGGCTGTCTCCAACCTCC 61

Qy 21673 TGACCTCATGATCTGCCACCTTGCCCTCCCAAAAGTGTGGGAT 21716

Db 62 TGACCTTGATCCACCAGCCTCGGCCCTCCCAAAAGTGTGGGAT 105

RESULT 4

US-08-485-862B-65

; Sequence 65, Application US/08485862B

; Patent No. 5989838

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court


```
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,862B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-0727
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-485-862B-65

Query Match 0.3%; Score 78.4; DB 4; Length 105;
Best Local Similarity 84.6%; Pred. No. 2.4e-08;
Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 21613 TTTTGTATTTTGTAGTAAGACGGGTTTCACCATGTTGGTCAGGCTGCTCTCCAACTCC 21672
      ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 TTTTACATCTTTAGTAGACAGGTTTCACCATATTTGGCCAGGCTGCTCTCAAACTCC 61

QY 21673 TGACCTCATGATCTGCCACCTTGGCCCTCCCAAGTCTGGGAT 21716
      ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 TGACCTTGTGATCCACGACCTCGGCCCTCCCAAGTCTGGGAT 105

RESULT 5
US-08-787-739-65
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-0334
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-787-739-65

Query Match 0.3%; Score 78.4; DB 5; Length 105;
Best Local Similarity 84.6%; Pred. No. 2.4e-08;
Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 21613 TTTTGTATTTTGTAGTAAGACGGGTTTCACCATGTTGGTCAGGCTGCTCTCCAACTCC 21672
      ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 TTTTACATCTTTAGTAGACAGGTTTCACCATATTTGGCCAGGCTGCTCTCAAACTCC 61

QY 21673 TGACCTCATGATCTGCCACCTTGGCCCTCCCAAGTCTGGGAT 21716
      ||||| ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db 62 TGACCTTGTGATCCACGACCTCGGCCCTCCCAAGTCTGGGAT 105

RESULT 6
US-08-481-658B-65/C
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 42A
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-481-658B-65

Query Match          0.2%; Score 68; DB 4; Length 1
Best Local Similarity 84.0%; Pred.No. 4.3e+06;
Matches      89; Conservative    0; Mismatches   15; Indel

QY 15260 ATACGAGCACTTTGGGAGGCCGATGTGGTGGATCACCTGAGTCAGGAGG
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 105 ATCCAGCACTTTGGGAGGCCGAGGTGGTGGATCACC--AAGGTCAGGAGG

QY 15320 GACTGGCCAAACATGGTGAAACCTCATCTCTAGTAAAAATACAAAA 153
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 47 GCCTGGCCAATATGGTGAACACCTGCTCTCTACTPAAAAGATGTAAAAA 2

```

```

RESULT 7
US-08-477-504A-65/c
; Sequence 65, Application US/08477504A
; Patent No. 5972353
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,504A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:

```

```
; NAME: Lauder, Leona L.  
; REGISTRATION NUMBER: 30,863  
; REFERENCE/DOCKET NUMBER: D-0021.3D  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-435-2034  
; TELEX: 415-435-2034  
; TELEFAX: 415-435-0727  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 105 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-477-504A-65
```

Query Match 0.2%; Score 68; DB 4; Length 105;
Best Local Similarity 84.0%; Pred.No. 4.3e-06;
Matches 89; Conservative 0; Mismatches 15; Indels 2; Gaps 1;

QY 15260 ATACCAGCACTTTGGGAGGCCGTATGGTGGTGACCCACTGTAGTAGTAATAAACAAAA 15319
 ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 105 ATCCAGCAGCACTTTGGGAGGCCGTATGGTGGTGAC--AAGGTCAGGAGTTTGAGAGCA 48
 ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 15320 GACTGGCGACAATGCTGAACCTCATCTCTAGTAAAATACAAAAA 15365
 ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 47 GCCTGGCCAATATGCTGAACCTGTCTCTACTAAAGATGTA AAAA 2

```
RESULT 8  
US-08-486-756A-65/c  
; Sequence 65, Application US/08486756A  
; Patent No. 5981711  
; GENERAL INFORMATION:  
; APPLICANT: Zavada, Jan  
; APPLICANT: Pastorekova, Silvia  
; APPLICANT: Pastorek, Jaromir  
; TITLE OF INVENTION: MN Gene and Protein  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Leona L. Lauder  
; STREET: 6 Mariposa Court  
; CITY: Tiburon  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94920  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,756A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/260,190  
; FILING DATE: 15-JUN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lauder, Leona L.  
; REGISTRATION NUMBER: 30,863  
; REFERENCE/DOCKET NUMBER: D-0021.3C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-435-2034  
; TELEFAX: 415-435-0727  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 105 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear
```

MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-486-756A-65

Query Match 0.2%; Score 68; DB 4; Length 105;
Best Local Similarity 84.0%; Pred. No. 4.3e-06;
Matches 89; Conservative 0; Mismatches 15; Indels 2; Gaps 1;

QY 15260 ATACGAGCACTTTGGAGGCCGATGGGTGGATCACCTGAGGTGAGGAGTTTGAGACCA 15319
|| |||||
DB 105 ATCCGAGCACTTTGGAGGCCGAGGCTGGTGGATCAC--AAGTCAAGGAGTTTGAGAGCA 48
|| |||||
QY 15320 GACTGCCAACATGGTGAACCTCTCTCTAGTAAAAATACAAAA 15365
|| |||||
DB 47 GCCTGCCAATATGGTGAACCCCTGCTCTACTAAAGATGTAATAA 2
|| |||||

RESULT 9
US-08-485-862B-65/c
; Sequence 65, Application US/08485862B
; Patent No. 5989838
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,862B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-485-862B-65

Query Match 0.2%; Score 68; DB 4; Length 105;
Best Local Similarity 84.0%; Pred. No. 4.3e-06;
Matches 89; Conservative 0; Mismatches 15; Indels 2; Gaps 1;

QY 15260 ATACGAGCACTTTGGAGGCCGATGGGTGGATCACCTGAGGTGAGGAGTTTGAGACCA 15319
|| |||||
DB 105 ATCCGAGCACTTTGGAGGCCGAGGCTGGTGGATCAC--AAGTCAAGGAGTTTGAGAGCA 48
|| |||||
QY 15320 GACTGCCAACATGGTGAACCTCTCTCTAGTAAAAATACAAAA 15365
|| |||||
DB 47 GCCTGCCAATATGGTGAACCCCTGCTCTACTAAAGATGTAATAA 2
|| |||||

RESULT 10
US-08-787-739-65/c
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-787-739-65

	;	TOPOLOGY:	both	
PCT-US95-17111A-70				
		Query Match	0.2%; Score 60.4; DB 6; Length 78;	
		Best Local Similarity	85.9%; Pred. No. 0.00016;	
		Matches	67; Conservative 0; Mismatches 11; Indels 0; Gaps 0;	
Oy	21598	ACCAAGCCCGGCTAAATTTTGTATTATTTAGTAAGAAGGGGTTTCACCATGTTGGTCAGG	21657	
Db	1	ACAAGCCCCAGCTAATTTGTATTATTTAGTAGAGATGGGGTTTCTCCATGTCATCAGG	60	
Oy	21658	CTGGTCTCCAACCTCCTGA	21675	
Db	61	CTGGTGCGAACTCCTGA	78	
RESULT	15			
US-08-454-557C-70/c		; Sequence 70, Application US/08454557C		
		; Patent No. 5830670		
		GENERAL INFORMATION:		
		APPLICANT: de la Monte, Suzanne		
		APPLICANT: Wands, Jack R.		
		TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection		
		TITLE OF INVENTION: of Alzheimer's Disease		
		NUMBER OF SEQUENCES: 121		
		CORRESPONDENCE ADDRESS:		
		ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.		
		STREET: 1100 New York Avenue, Suite 600		
		CITY: Washington		
		STATE: D.C.		
		COUNTRY: U.S.A.		
		ZIP: 20005-3934		
		COMPUTER READABLE FORM:		
		MEDIUM TYPE: Floppy disk		
		COMPUTER: IBM PC compatible		
		OPERATING SYSTEM: PC-DOS/MS-DOS		
		SOFTWARE: PatentIn Release #1.0, Version #1.25		
		CURRENT APPLICATION DATA:		
		APPLICATION NUMBER: US/08/454,557C		
		FILING DATE: 30-MAY-1995		
		CLASSIFICATION: 514		
		ATTORNEY/AGENT INFORMATION:		
		NAME: Ludwig, Steven R.		
		REGISTRATION NUMBER: 36,203		
		REFERENCE/DOCKET NUMBER: 0609.3840003		
		TELECOMMUNICATION INFORMATION:		
		TELEPHONE: (202) 371-2600		
		TELEFAX: (202) 371-2540		
		INFORMATION FOR SEQ ID NO: 70:		
		SEQUENCE CHARACTERISTICS:		
		LENGTH: 78 base pairs		
		TYPE: nucleic acid		
		STRANDEDNESS: both		
		TOPOLOGY: both		
US-08-450-673C-70				
		Query Match	0.2%; Score 60.4; DB 4; Length 78;	
		Best Local Similarity	85.9%; Pred. No. 0.00016;	
		Matches	67; Conservative 0; Mismatches 11; Indels 0; Gaps 0;	
Oy	21598	ACCAAGCCCGGCTAAATTTTGTATTATTTAGTAAGAAGGGGTTTCACCATGTTGGTCAGG	21657	
Db	1	ACAAGCCCCAGCTAATTTGTATTATTTAGTAGAGATGGGGTTTCTCCATGTTTCATCAGG	60	
Oy	21658	CTGGTCTCCAACCTCCTGA	21675	
Db	61	CTGGTGCGAACTCCTGA	78	
RESULT	14			
PCT-US95-17111A-70		; Sequence 70, Application PC/TUS9517111A		
		GENERAL INFORMATION:		
		APPLICANT: de la Monte, Suzanne		
		APPLICANT: Wands, Jack R.		
		TITLE OF INVENTION: Neural Thread Protein Gene Expression and		
		TITLE OF INVENTION: Detection of Alzheimer's Disease		
		NUMBER OF SEQUENCES: 121		
		CORRESPONDENCE ADDRESS:		
		ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.		
		STREET: 1100 New York Avenue, Suite 600		
		CITY: Washington		
		STATE: D.C.		
		COUNTRY: U.S.A.		
		ZIP: 20005-3934		
		COMPUTER READABLE FORM:		
		MEDIUM TYPE: Floppy disk		
		COMPUTER: IBM PC compatible		
		OPERATING SYSTEM: PC-DOS/MS-DOS		
		SOFTWARE: PatentIn Release #1.0, Version #1.25		
		CURRENT APPLICATION DATA:		
		APPLICATION NUMBER: PCT/US95/17111A		
		FILING DATE:		
		CLASSIFICATION:		
		PRIOR APPLICATION DATA:		
		APPLICATION NUMBER: 08/340,426		
		FILING DATE: 14-NOV-1994		
		ATTORNEY/AGENT INFORMATION:		
		NAME: Ludwig, Steven R.		
		REGISTRATION NUMBER: 36,203		
		REFERENCE/DOCKET NUMBER: 0609.3840002		
		TELECOMMUNICATION INFORMATION:		
		TELEPHONE: (202) 371-2600		
		TELEFAX: (202) 371-2540		
		INFORMATION FOR SEQ ID NO: 70:		
		SEQUENCE CHARACTERISTICS:		
		LENGTH: 78 base pairs		
		TYPE: nucleic acid		
		STRANDEDNESS: both		

Search completed: June 19, 2000, 11:34:48
Job time: 427218 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2000, 11:28:29 ; Search time 17971.8 Seconds
(without alignments)
-1569.789 Million cell updates/sec

Title: US-08-852-495C-2_COPY_140000_169000
Perfect score: 29001
Sequence: 1 GGTTCGACAAAGGTGTCAA.....TCCTTCAGGAGTTACTTCTA 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_bal.*
2: gb_ba2.*
3: gb_om.*
4: gb_ov.*
5: gb_pat.*
6: gb_ph.*
7: gb_pl1.*
8: gb_pl2.*
9: gb_pr1.*
10: gb_pr2.*
11: gb_pr3.*
12: gb_ro.*
13: gb_sts.*
14: gb_sy.*
15: gb_un.*
16: gb_v1.*
17: em_fun.*
18: em_hum1.*
19: em_hum2.*
20: em_in.*
21: em_om.*
22: em_or.*
23: em_ov.*
24: em_pat.*
25: em_ph.*
26: em_pl.*
27: em_ro.*
28: em_sts.*
29: em_sy.*
30: em_un.*
31: em_v1.*
32: gb_htg1.*
33: gb_htg2.*
34: gb_in1.*
35: gb_in2.*
36: em_bal.*
37: em_ba2.*
38: em_hum3.*
39: em_hum4.*
40: gb_pr4.*
41: gb_htg3.*
42: gb_htg4.*
43: gb_htg5.*
44: gb_htg6.*

45: gb_htg7.*
46: em_htg1.*
47: em_htg2.*
48: em_htg3.*
49: em_hum5.*
50: gb_pl3.*
51: gb_pr5.*
52: gb_htg8.*
53: gb_htg9.*
54: gb_htg10.*
55: gb_htg11.*
56: gb_htg12.*
57: gb_htg13.*
58: gb_htg14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	85.8	0.3	108	11	HSU67803 U67803 Human small
2	85.2	0.3	108	10	HSLDLRN2 X05250 Human LDL-r
3	80.6	0.3	107	9	HUMALCE162 M87924 Human carc1
4	79.2	0.3	104	9	HUMALCE272 M87899 Human carc1
5	78.2	0.3	103	9	HUMALCE221 M87896 Human carc1
c 6	77.6	0.3	108	10	HSLDLRN2 X05250 Human LDL-r
7	77.8	0.3	108	11	HSU67804 U67804 Human small
8	76.2	0.3	108	11	HSU67808 U67808 Human small
9	75.4	0.3	103	13	HS8IC8R X57789 Human seque
c 10	75	0.3	97	9	HUMLDLRA1 M14178 Human low d
c 11	75	0.3	107	9	HUMALCE162 M87924 Human carc1
c 12	74.4	0.3	103	13	HS8IC8R X57789 Human seque
c 13	74.6	0.3	108	10	HSLDLRD1 X05249 Human LDL-r
14	74.6	0.3	108	10	HSLDLRD2 X05251 Human LDL-r
15	74.4	0.3	110	11	HSU67807 U67807 Human small
16	74	0.3	91	13	HUMUT8164A L30244 Human STS U
c 17	72.4	0.2	101	10	S79560 HRX (intron
c 18	72.4	0.2	103	9	HUMALCE221 M87896 Human carc1
c 19	72.4	0.2	108	10	HSLDLI12 X05248 Human LDL-r
c 20	72.6	0.3	108	13	G43535 WIAF-2393-S
21	72.4	0.2	110	9	HUMALCE43 M87900 Human carc1
c 22	72	0.2	97	9	HUMLDLRA2 M14180 Human low d
c 23	71.8	0.2	110	11	HSU67807 U67807 Human small
c 24	71.4	0.2	97	9	HUMLDLRDJ M14179 Human fam1
25	71	0.2	97	9	HUMLDLRA2 M14180 Human low d
c 26	70.8	0.2	100	13	HUMUT931A L31299 Human STS U
c 27	71	0.2	105	13	G32655 G32655 A009L30 Hum
28	70.2	0.2	97	9	HUMLDLRA1 M14178 Human low d
29	69.8	0.2	107	11	HSU67806 U67806 Human small
30	69.4	0.2	108	9	HUMDI03M5 D16965 Human HepG2
c 31	68.4	0.2	95	13	HUMUT8002B L30176 Human STS U
c 32	68.6	0.2	97	9	HUMLDLRDJ M14179 Human fam1
33	68.6	0.2	108	10	HSLDLRD1 X05249 Human LDL-r
c 34	68.6	0.2	108	10	HSLDLRD2 X05251 Human LDL-r
35	67.6	0.2	90	9	HUMLDLRM M15365 Human low d
36	67.8	0.2	108	13	G43535 G43535 WIAF-2393-S
37	67.2	0.2	100	9	HUMGALNSA D45223 Human GALNS
c 38	66.8	0.2	95	10	HSSTHPKIB X66361 H. sapiens m
39	66	0.2	100	10	HSLAS27 X91545 H. sapiens D
40	66	0.2	100	13	HUMUT931A L31299 Human STS U
c 41	65.6	0.2	90	9	HUMLDLRFL K03555 Human low d
42	65.8	0.2	101	10	S79560 HRX (intron
c 43	65.4	0.2	84	5	AR051521 AR051521 Sequence
c 44	65.6	0.2	102	13	G32906 G32906 A009W09 Hum
45	65.6	0.2	109	13	G02323 G02323 human STS S

ALIGNMENTS

```
RESULT 1
HSU67803      108 bp      RNA      PRI      01-AUG-1997
LOCUS         Human small cytoplasmic ALU transcript.
DEFINITION    U67803
ACCESSION     U67803.1 GI:2289917
VERSION       U67803.1
KEYWORDS      ALu.
SOURCE        human.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
               Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE     1 (bases 1 to 108)
AUTHORS       Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE         cDNAs derived from primary and small cytoplasmic ALu (sALu)
               transcripts
JOURNAL       J. Mol. Biol. 271 (2), 222-234 (1997)
MEDLINE       97415756
REFERENCE     2 (bases 1 to 108)
AUTHORS       Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE         Direct Submission
JOURNAL       Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
               Children's Hospital of Philadelphia, 1004F Abramson Research
               Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
               Location/Qualifiers
FEATURES      1..108
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="fscAlu2"
               /note="sAlu"
               /rpt_family="Alu"
               /rpt_type="dispersed"
               /rpt_type="dispersed"
BASE COUNT    23 a 39 c 30 g 16 t
ORIGIN
               repeat_region
               1..108
               /note="sAlu"
               /db_xref="taxon:9606"
               /clone="fscAlu2"
               1..108
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /cell_line="Ntera2D1"
               /dev_stage="embryo"
               /sex="male"
               /tissue_type="carcinoma"
Query Match    0.3%; Score 85.6; DB 11; Length 108;
Best Local Similarity 92.8%; Pred. No. 0.0008;
Matches 90; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 26228 GCCTGTATCCACGACTTTGGGAGGCTGGTGGTGAATCAGAGGTGAGGATCAAA 26287
Db 1 GCCTGTATCCACGACTTTGGGAGGCTGGTGGTGAATCAGAGGTGAGGATCAAA 60
Qy 26288 GACCATCTGCGCCACATGTTGAACCCCGTCTCTAC 26324
Db 61 GACCATCTGCGCTAACAGGTGAACCCCGTCTCTAC 97

RESULT 2
HSLDLN2      108 bp      DNA      PRI      20-MAY-1992
LOCUS         Human LDL-receptor gene Intron 14 fragment (normal gene).
DEFINITION    X05250
ACCESSION     X05250
VERSION       X05250.1 GI:34337
KEYWORDS      Alu repetitive sequence; low density lipoprotein receptor.
SOURCE        human.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
               Primates; Catarrhini; Homiidae; Homo.
REFERENCE     1 (bases 1 to 108)
AUTHORS       Horsthemke,B., Beisiegel,U., Dunning,A., Hovinga,J.R.,
               Williamson,R. and Humphries,S.
TITLE         Unequal crossing-over between two alu-repetitive DNA sequences in
               the low-density-lipoprotein-receptor gene. A possible mechanism for
               the defect in a patient with familial hypercholesterolaemia
JOURNAL       Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE       87161901
COMMENT       See X05252 for deletion junction
               Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
               Location/Qualifiers
FEATURES      1..108
```

```
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                1..108
                /note="intron XIV fragment"
BASE COUNT    28 a 23 c 39 g 18 t
ORIGIN
               Query Match    0.3%; Score 85.2; DB 10; Length 108;
               Best Local Similarity 87.7%; Pred. No. 0.00098;
               Matches 93; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Qy 26334 AAAAATAGCCAGCATGGTAGCACATGCTGTATCCAGCTACTCAAGAGGCTGAGGC 26393
Db 3 AAAAATAGCCAGCGTGGTGGCAGGTGCTGTATCCAGCTACTCGGAGGCTGAGGC 62
Qy 26394 AGGGGAATTGCTTAACCCGGGAGGTGGACATTCAGTGAGCTGAG 26439
Db 63 AGGAGAATTGCTTAACCCAGGAGGAGGAGGTTCAGTGAGCCGAG 108

RESULT 3
HUMALCE162   107 bp ss-RNA  PRI      15-APR-1994
LOCUS         Human carcinoma cell-derived Alu RNA transcript, clone CE162.
DEFINITION    M87924
ACCESSION     M87924.1 GI:174871
KEYWORDS      Alu repeat.
SOURCE        Homo sapiens male embryo carcinoma cDNA to other RNA.
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
               Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE     1 (bases 1 to 107)
AUTHORS       Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE         Alu RNA transcripts in human embryonal carcinoma cells. Model of
               post-transcriptional selection of master sequences
JOURNAL       J. Mol. Biol. (1992) In press
FEATURES      1..107
               Location/Qualifiers
               source
               1..107
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /cell_line="Ntera2D1"
               /dev_stage="embryo"
               /sex="male"
               /tissue_type="carcinoma"
BASE COUNT    28 a 30 c 35 g 14 t
ORIGIN
               Query Match    0.3%; Score 80.6; DB 9; Length 107;
               Best Local Similarity 86.4%; Pred. No. 0.0047;
               Matches 89; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 2906 GCGAGGAGAAATCACTTGAACCTGGGAGGAGAGGTTGAGTGAGGGAGGAGTGGGCCACT 2965
Db 5 GCGAGGAGAAATGCGGTGAACCCGGGAGGAGGAGTTGAGTGAGCGGAGATCGGCCACT 64
Qy 2966 GCACCTCCAGCTGAGCACACAGCGAGACTCTGTCTCAAAAAA 3008
Db 65 GCACCTCCAGCTGCGGCGACAGCGAGACTCCGTCTCAAAAAA 107

RESULT 4
HUMALCE272   104 bp ss-RNA  PRI      15-APR-1994
LOCUS         Human carcinoma cell-derived Alu RNA transcript, clone CE272.
DEFINITION    M87899
ACCESSION     M87899.1 GI:174875
KEYWORDS      Alu repeat.
SOURCE        Homo sapiens male embryo carcinoma cDNA to other RNA.
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
               Eutheria; Primates; Catarrhini; Homiidae; Homo.
```



```

LOCUS       HSLDRN2      108 bp      DNA                     PRI      20-MAY-1992
DEFINITION   Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION    X05250
VERSION      X05250.1  GI:34337
KEYWORDS     Alu repetitive sequence; low density lipoprotein receptor.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
              Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1  (bases 1 to 108)
AUTHORS      Horstchenke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
              Williamson,R. and Humphries,S.
TITLE        Unequal crossing-over between two alu-repetitive DNA sequences in
              the low-density-lipoprotein-receptor gene. A possible mechanism for
              the defect in a patient with familial hypercholesterolaemia
JOURNAL      Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE      87161901
COMMENT      See X05252 for deletion junction
              Data kindly received (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1..108
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
     intron          1..108
                    /note="intron XIV fragment"
BASE COUNT        28 a   23 c   39 g   18 t
ORIGIN
Query Match           0.3%; Score 77.6; DB 10; Length 108;
Best Local Similarity 82.4%; Pred. No. 0.013;
Matches 89; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 13383 CTCAACTCAGTCGAACCTCCACCCTCGGATTACGGCAATTTCTCTGCCTCAGCCTCCGG 13442
|||||
Db 108 CTGGCTCATCTGCAACTCTGCCCTCTCGGGTTCAGCAATTTCTCTGCCTCAGCCTCCGG 49
|||||

QY 13443 AGTAACCTCGCACTACAGTGCGCACACCACCAAGTGCTAAATTTTTT 13490
|||||
Db 48 AGTAGCTGGGATTACAGCAGCAGCTGCCACGCGCTGGCTAATTTTGT 1

RESULT 7
HSU67804
LOCUS       HSU67804      108 bp      RNA                     PRI      01-AUG-1997
DEFINITION   Human small cytoplasmic Alu transcript.
ACCESSION    U57804
VERSION      U67804.1  GI:2289918
KEYWORDS     Alu.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Mammalia;
              Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1  (bases 1 to 108)
AUTHORS      Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE        cDNAs derived from primary and small cytoplasmic Alu (scAlu)
              transcripts
JOURNAL      J. Mol. Biol. 271 (2), 223-234 (1997)
MEDLINE      97415756
REFERENCE    2  (bases 1 to 108)
AUTHORS      Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE        Direct Submission
JOURNAL      Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
              Children's Hospital of Philadelphia, 1004F Abramson Research
              Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES             Location/Qualifiers
     source          1..108
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="TscAlu3"
     repeat_region   1..108
                    /note="scAlu"
                    /rpt_family="Alu"

```

```

BASE COUNT      26 a      38 c      26 g      18 t
ORIGIN

Query Match      0.3%; Score 77.8; DB 11; Length 108;
Best Local Similarity 87.6%; Pred. No. 0.012;
Matches 85; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 26228 GCCTGTATCCACACCTTTGGGAGGCTGAGGTGGTGAATCAGCAGGTCTAGGAGATCAA 26287
|||||
Db 1 GCCTGTATCCACACCTTTGGGAGGCTGAGGTGGTGAATCAGCAGGTCTAGGAGATCAA 60
|||||

QY 26288 GACCATCTGCCCAACATGGTGAACCCCGCTCTCTAC 26324
|||||
Db 61 GACCATCTGCCCAACATGGTGAACCCCGCTCTTTC 97
|||||

RESULT 8
HSU67808      108 bp      RNA      PRI      01-AUG-1997
LOCUS      Human small cytoplasmic Alu transcript.
ACCESSION      U67808
VERSION      U67808.1 GI:2289922
KEYWORDS      Alu.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 108)
AUTHORS      Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE      cDNAs derived from primary and small cytoplasmic Alu (scAlu)
JOURNAL      J. Mol. Biol. 271 (2), 222-234 (1997)
MEDLINE      97413756
AUTHORS      Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE      Direct Submission
JOURNAL      Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
Children's Hospital of Philadelphia, 1004F Abramson Research
Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES      Location/Qualifiers
source      1..108
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="TscAlu7"
repeat_region 1..108
            /notes="scAlu"
            /rpt_family="Alu"
            /rpt_type="dispersed"
BASE COUNT      22 a      37 c      28 g      21 t
ORIGIN

Query Match      0.3%; Score 76.2; DB 11; Length 108;
Best Local Similarity 86.6%; Pred. No. 0.021;
Matches 84; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 26228 GCCTGTATCCACACCTTTGGGAGGCTGAGGTGGTGAATCAGCAGGTCTAGGAGATCAA 26287
|||||
Db 1 GCCTGTATCCACACCTTTGGGAGGCTGAGGTGGTGAATCAGCAGGTCTAGGAGATTTGA 60
|||||

QY 26288 GACCATCTGCCCAACATGGTGAACCCCGCTCTCTAC 26324
|||||
Db 61 GACCAGCTGCCCAACATGGTGAACCTCCGCTCTTTC 97
|||||

RESULT 9
HS81C8R      103 bp      DNA      STS      05-SEP-1991
LOCUS      Human sequence tagged site 81C8R DNA from 19q13.
DEFINITION      X57789
ACCESSION      X57789
VERSION      X57789.1 GI:23938

KEYWORDS      STS; myotonic dystrophy.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 103)
AUTHORS      Aldridge,F.L.
TITLE      Direct Submission
JOURNAL      Alderley Park, Macclesfield, Cheshire, SK10 4TG, UK
2 (bases 1 to 103)
AUTHORS      Butler,R., Riley,J.H., Ogilvie,D.J., Anand,R., Buxton,J.,
Davies,J., Johnson,K. and Markham,A.F.
TITLE      Two sequence-tagged sites defining the ends of a 380 kb YAC clone
from 19q13
JOURNAL      Nucleic Acids Res. 19 (17), 4787 (1991)
MEDLINE      91367697
COMMENT      See also X57788 for STS 81C8L.
FEATURES      Location/Qualifiers
source      1..103
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="19q13"
            /germline
            /clone_lib="YAC library: ICI"
            /clone="81C8"
BASE COUNT      29 a      28 c      23 g      22 t      1 others
ORIGIN

Query Match      0.3%; Score 75.4; DB 13; Length 103;
Best Local Similarity 87.2%; Pred. No. 0.028;
Matches 82; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 13498 CTGTAGAGATGGGTTTCGCCAGTGTGGCCAGGTGTCTCAAACTCTGACCTGAAGTG 13557
|||||
Db 1 CAGTAGAGATGGGTTTCACCATTGTGGCCAGGTGTCTCAAGAACTCTGACCTTAAGTG 60
|||||

QY 13558 TTCCACCCACCTCGCGCTCCCAAGTGTCTGATT 13591
|||||
Db 61 ATCCACCCACCTCGACCTCCCAAGTGCNGAAT 94
|||||

RESULT 10
HUMDLRAL/C      97 bp      DNA      PRI      07-JAN-1995
LOCUS      Human low density lipoprotein receptor gene, intron 4 (partial).
DEFINITION      M14178
ACCESSION      M14178.1 GI:187097
KEYWORDS      low density lipoprotein receptor-1.
SEGMENT      1 of 2
SOURCE      Human white blood cell DNA.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 97)
AUTHORS      Hobbs,H.H., Brown,M.S., Goldstein,J.L. and Russell,D.W.
TITLE      Deletion of exon encoding cysteine-rich repeat of low density
lipoprotein receptor alters its binding specificity in a subject
with familial hypercholesterolemia
JOURNAL      J. Biol. Chem. 261 (28), 13114-13120 (1986)
MEDLINE      87008518
COMMENT      Analysis of the LDL-receptor gene of a patient with familial
hypercholesterolemia (FH) revealed the deletion of exon 5 resulting
from a homologous recombination between repetitive Alu sequences of
intron 4 and intron 5.
FEATURES      Location/Qualifiers
source      1..97
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /map="19p13.3"
            <1..>97
            /gene="LDLR"
intron

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2000, 11:47:50 ; Search time 580.8 seconds
(without alignments)
12492.804 Million cell updates/sec

Title: US-08-852-495C-2_COPY_140000_169000
Perfect score: 29001
Sequence: 1 GGTTCGACAAAGGTCTCAA.....TCCTTCAGGAGTACTTCTCA 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 433070

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	80.4	0.3	108	1 T25009	Human gene signatu
C 2	73.2	0.3	108	1 X12095	Human biallelic po
C 3	68.8	0.2	100	1 T24892	Human gene signatu
C 4	68.4	0.2	108	1 X12095	Human biallelic po
C 5	63.4	0.2	100	1 X12087	Human biallelic po
C 6	63.4	0.2	100	1 X12085	Human biallelic po
C 7	62.4	0.2	100	1 T24892	Human gene signatu
C 8	62.2	0.2	100	1 X12086	Human biallelic po
C 9	62.2	0.2	106	1 V11611	Human biallelic po
C 10	61.2	0.2	91	1 T25854	Homo sapiens adult
C 11	61.2	0.2	103	1 T20927	Human gene signatu
C 12	60.6	0.2	101	1 V00420	3' fragment of clo
C 13	60.6	0.2	103	1 T26213	Human gene signatu
C 14	60.2	0.2	92	1 V11595	Homo sapiens adult
C 15	60.2	0.2	92	1 V61480	Human secreted pro
C 16	60.4	0.2	108	1 T26828	Human gene signatu
C 17	59.6	0.2	84	1 V04275	Secreted protein C
C 18	59.6	0.2	84	1 V09269	Nucleotide sequenc
C 19	59.4	0.2	85	1 V11441	Human secreted pro
C 20	59.4	0.2	87	1 V05740	Nucleotide sequenc
C 21	59.4	0.2	94	1 V05738	Nucleotide sequenc
C 22	59.2	0.2	99	1 T91300	Human M97-2 secret
C 23	59.4	0.2	108	1 T25009	Human gene signatu
C 24	58.4	0.2	69	1 Q29016	Probe to internal
C 25	58.4	0.2	79	1 V26706	Human novel secret
C 26	58.6	0.2	85	1 V05720	Nucleotide sequenc
C 27	58.4	0.2	106	1 V00430	3' fragment of clo
C 28	57.4	0.2	77	1 V77194	Staphylococcus aur
C 29	57.4	0.2	80	1 V30930	Human secreted pro
C 30	57.4	0.2	93	1 T22572	Human gene signatu
C 31	56.8	0.2	69	1 V21236	Homo sapiens clone
C 32	56.8	0.2	70	1 V30923	Human secreted pro
C 33	56.8	0.2	74	1 V99725	Human adult testis
C 34	56.2	0.2	103	1 T26213	Human gene signatu

35	55.8	0.2	69	1 V21226	Homo sapiens clone
C 36	55.8	0.2	93	1 T24259	Human gene signatu
C 37	55.2	0.2	87	1 T21566	Human gene signatu
C 38	55.2	0.2	87	1 T21566	Human gene signatu
C 39	55.2	0.2	97	1 T26728	Human gene signatu
C 40	55.4	0.2	100	1 X12087	Human biallelic po
C 41	55.4	0.2	100	1 X12085	Human biallelic po
C 42	55.4	0.2	100	1 X12086	Human biallelic po
C 43	54.8	0.2	99	1 T20931	Human gene signatu
C 44	54	0.2	95	1 T23131	Human gene signatu
C 45	53.8	0.2	109	1 T23895	Human gene signatu

ALIGNMENTS

RESULT 1
T25009/c
ID T25009 standard; cDNA to mRNA; 108 BP.
AC T25009;
DT 07-NOV-1996 (first entry)
DE Human gene signature HUMG07131.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
human; cloning; mapping; non-biased library; diagnosis; detection;
cell typing; abnormal cell function; ss.
KW Homo sapiens.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATSU) MATSUBARA K.
PA (OKUBU) OKUBO K.
PI Matsubara K. Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 1748; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 34 A; 31 C; 26 G; 15 T;

Query Match 0.3%; Score 80.4; DB 1; Length 108;
Best Local Similarity 83.3%; Pred. No. 0.0064;
Matches 90; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
Oy 9891 TTTTCTTTTCTTTTCTTTGAGACAAGGCTTGTCTCTGCTAGCTTGGATTCAGTG 9950
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 108 TTGTTGTGTGTGTGTGTGTTTCAACAGGGTCTTGTCTCTGCTAGCTTGGATTCAGTG 49
Oy 9951 GCATGACCATGCTCAGCTGCAGCCCTTGCCTCTAGGCTCAAGCAATC 9998
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 48 GCGTGACCATGCTCAGCTGCAGCCCTTGCCTCTAGGCTCAAGCAATC 1
RESULT 2
X12095/c
ID X12095 standard; DNA; 108 BP.


```
PR 12-NOV-1993; JP-355504.
PA (MATSU) MATSUBARA K.
PA (OKUBU) OKUBO K.
PI Matsubara K, Okubo K;
PI WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 1720: 2245pp: Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues: synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;

Query Match 0.2%; Score 62.4; DB 1; Length 100;
Best Local Similarity 75.8%; Pred. No. 1.5;
Matches 75; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 18746 TTTTGTGTTGAGATGGAGTCTTACTCTGCTGCTCAAGCTGGAGTGCAGTGCACCAATC 18805
DB 100 TTTGTTTTCACACAGAGTGTACCTCTGTCCACCGGCGAGTGCAGNGTGCATC 41

QY 18806 TCAGCTCACTGCAACCTCTGCTCTCTCTGGGTTCAAGCAAT 18844
DB 40 TCAGCTNATTGCAAAATCTGCTCTCCAGGTTCAAGCGAT 2

RESULT 8
X12086/c
ID X12086 standard; DNA; 100 BP.
AC X12086;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment EST98276b.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medication;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
PI WPI: 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
```

```
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match 0.2%; Score 62.2; DB 1; Length 100;
Best Local Similarity 76.8%; Pred. No. 1.5;
Matches 76; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 5818 GTGCTCAAGCTCTAATCCCAACACTTTGGGAGGCTAAGTGGGAGGATGCTTGAGCC 5877
DB 99 GTGACTCACACCTATAATCTGGCAGCTTTAGGAGGCTKAGGAAGGAGGATTTGTTGAAC 40

QY 5878 CAGTAGTTCAAGACGACGCTGGGCAACATGAGAGAACCC 5916
DB 39 CAGGAGCTCAGACCATCTCTGGGAAACATAGCAAGACTC 1

RESULT 9
V11611
ID V11611 standard; cDNA; 106 BP.
AC V11611;
DT 11-SEP-1998 (first entry)
DE Homo sapiens adult retina clone BO365_2 3' region.
KW adult; retina; cDNA library; clone BO365_2; anti-inflammatory;
KW therapeutic composition; autoimmune disease; immune; stimulation;
KW suppression; ds.
OS Homo sapiens.
PN WO9814576-A2.
PD 09-APR-1998.
PF 03-OCT-1997; U18007.
PR 04-OCT-1996; US-726237.
PA (GEMY ) GENETICS INST INC.
PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
PI Racie LA, Spaulding V, Treacy M;
PI WPI: 98-240082/21.
PT Nucleic acids encoding novel secreted proteins - useful as, e.g.
PT anti-inflammatory, immuno-stimulatory or suppressing agents
PS Claim 30; Page 73; 110pp; English.
CC The sequence is that of an isolated polynucleotide which may
CC be of use in the production of therapeutic compositions for
CC treating or ameliorating a medical condition in a mammal. Such
CC compositions may be used for, e.g. research purposes as markers for
CC tissues, molecular weight markers for gels, primers or probes, for
CC nutrition as carbon, nitrogen or carbohydrate source. They can also be
CC used as a cytokine for cell proliferation and differentiation activity,
CC as immune stimulants or suppressors, e.g. for viral, bacterial or fungal
CC infections, for autoimmune diseases such as multiple sclerosis or
CC systemic lupus erythematosus, to regulate haematopoiesis, for tissue
CC growth, as an activator or inhibitor, or as a chemotactic or
CC chemokinetic, haemostatic and thrombocytic, receptor/ligand,
CC anti-inflammatory or tumour inhibitor agents.
SQ Sequence 106 BP; 104 A; 2 C; 0 G; 0 U;

Query Match 0.2%; Score 62.2; DB 1; Length 106;
Best Local Similarity 76.8%; Pred. No. 1.5;
Matches 76; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 7897 CAAAGAGAACCATCATCAGCCATGAAACACATGAAGGAAAGAGAAAGAGAAAGAGAA 7956
DB 6 CAAAGAGAACCATCATCAGCCATGAAACACATGAAGGAAAGAGAAAGAGAAAGAGAA 65

QY 7957 GAAAGAGAACCATCATCAGCCATGAAACACATGAAGGAAAGAGAAAGAGAA 7995
DB 66 AAGAGAGAACCATCATCAGCCATGAAACACATGAAGGAAAGAGAAAGAGAAAGAGAA 104
```

```
RESULT 10
T25854
AC T25854 standard; cDNA to mRNA; 91 BP.
AD T25854;
DE 22-OCT-1996 (first entry)
DE Human gene signature HUMGS08084.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1944; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 91 BP; 18 A; 22 C; 28 G; 18 T;

Query Match 0.2%; Score 61; DB 1; Length 91;
Best Local Similarity 78.7%; Pred. No. 2.2;
Matches 70; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 15897 ATCACTTGAGCGGGAGGCGAGAGGTTGACGTGAGCTGAGATTTGCGCCACTACACTACAG 15956
Dbb 2 ATCACTTGAGCGCTAGGAGCGAGGTTCAAGTGAGCTGAGATGCGCACTCTCTGCGCTCCAG 61

Qy 15957 CTGGGTGACAGAGAGATTTCTGTCTCA 15985
Dbb 62 CTTGGTGCACGCGTGAGANNCTGTCTCA 90

RESULT 11
T20927/c
AC T20927 standard; cDNA to mRNA; 103 BP.
AD T20927;
DE 24-JUL-1996 (first entry)
DE Human gene signature HUMGS02180.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
```

```
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 758-759; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 103 BP; 22 A; 27 C; 21 G; 31 T;

Query Match 0.2%; Score 61.2; DB 1; Length 103;
Best Local Similarity 75.0%; Pred. No. 2.1;
Matches 75; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 15800 CATGCCAAPACCCCTGCTCTACCAAAAATACAAATAGCTGGCAATGTGGCACATGC 15859
Dbb 100 CATGGAGAAATACTGCTCCCTACTNAAATATACNAAATCAGCTGGACATGGTGGCACACAC 41

Qy 15860 CTGTAAATCCCACTACTTGGGAGGCTGAAGCACACAGAATC 15899
Dbb 40 CTGTAGCCACAGCTACTTGGGAGGTGGAAGTGGGAGGATC 1

RESULT 12
V00420
ID V00420 standard; cDNA; 101 BP.
AC V00420;
DE 12-MAY-1998 (first entry)
DE 3' fragment of clone M97-2.
KW Human; secreted protein; molecular weight marker; genetic fingerprinting;
KW antibody production; nutritional supplement; therapy; neural tissue;
KW glioblastoma line TG986; clone M97-2; ds.
OS Homo sapiens.
PN W09740069-A2.
PD 30-OCT-1997.
PF 14-APR-1997; U06134.
PR 19-APR-1996; US-635311.
PA (GEM) GENETICS INST INC.
PI Jacobs K, Lavellie ER, McCoy JM, Merberg D, Racie LA,
PI Spaulding V;
DR WPI: 97-535776/49.
PT Isolated nucleic acid clones from ATCC 98028 encode novel secreted
PT proteins - having many potential uses, e.g. as immunomodulators,
PT cell proliferation or differentiation inhibitors or haematopoiesis
PT regulators
PS Claim 25; Page 70; 114pp; English.
CC This sequence represents the 3' end of clone M97-2, which is a
CC polynucleotide of the invention. This sequence was isolated from a human
CC neural tissue (glioblastoma line TG986) cDNA library. The polynucleotide,
CC which encodes a secreted protein, can be used, e.g. as a tissue or
CC molecular weight marker, in genetic fingerprinting, to raise anti-protein
CC or anti-DNA antibodies and in interaction trap assays. The protein can be
CC used to assay biological activity, raise antibodies for use in
CC immunoassays, as a marker, to identify inhibitors of its interactions and
CC as a nutritional supplement. It may also have a very wide range of
CC therapeutic and biological activities (no examples are given to support
CC this), e.g. cytokine or modulator of cell proliferation and
CC differentiation, immunostimulant or immunosuppressant, haematopoiesis
CC regulator, bone, cartilage, tendon, ligament and/or nerve tissue growth
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2000, 06:28:49 ; Search time 8512.45 Seconds
(without alignments)
13808.895 Million cell updates/sec

Title: US-08-852-495C-2_COPY_140000_169000
Perfect score: 29001
Sequence: 1 GGTTTTCACAAAGGTGTCAA.....TCCTTCAGAGTACTTCTA 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

- EST:*
- 1: em_est1:*
 - 2: em_est2:*
 - 3: em_est3:*
 - 4: em_est4:*
 - 5: em_est5:*
 - 6: em_est6:*
 - 7: em_est7:*
 - 8: em_est8:*
 - 9: em_est9:*
 - 10: em_est10:*
 - 11: em_est11:*
 - 12: em_est12:*
 - 13: em_est13:*
 - 14: em_est14:*
 - 15: em_est15:*
 - 16: em_est16:*
 - 17: em_est17:*
 - 18: em_est18:*
 - 19: em_est19:*
 - 20: gb_est1:*
 - 21: gb_est2:*
 - 22: gb_est3:*
 - 23: gb_est4:*
 - 24: gb_est5:*
 - 25: gb_est6:*
 - 26: gb_est7:*
 - 27: gb_est8:*
 - 28: gb_est9:*
 - 29: gb_est10:*
 - 30: gb_est11:*
 - 31: gb_est12:*
 - 32: gb_est13:*
 - 33: gb_est14:*
 - 34: gb_est15:*
 - 35: gb_est16:*
 - 36: gb_est17:*
 - 37: gb_est18:*
 - 38: gb_est19:*
 - 39: gb_est20:*
 - 40: gb_est21:*
 - 41: gb_est22:*
 - 42: gb_est23:*
 - 43: gb_est24:*
 - 44: gb_est25:*

- 45: gb_est26:*
- 46: gb_est27:*
- 47: gb_est28:*
- 48: gb_est29:*
- 49: gb_est30:*
- 50: gb_est31:*
- 51: gb_est32:*
- 52: em_est20:*
- 53: em_est21:*
- 54: em_est22:*
- 55: em_est23:*
- 56: em_est24:*
- 57: em_est25:*
- 58: em_est26:*
- 59: gb_est33:*
- 60: gb_est34:*
- 61: gb_est35:*
- 62: gb_est36:*
- 63: gb_est37:*
- 64: gb_est38:*
- 65: em_est27:*
- 66: em_est28:*
- 67: em_est29:*
- 68: em_est30:*
- 69: gb_est39:*
- 70: gb_est40:*
- 71: gb_est41:*
- 72: gb_est42:*
- 73: gb_est43:*
- 74: gb_est44:*
- 75: em_est31:*
- 76: em_est32:*
- 77: em_est33:*
- 78: em_est34:*
- 79: gb_est45:*
- 80: gb_est46:*
- 81: gb_est47:*
- 82: gb_gss1:*
- 83: gb_gss2:*
- 84: gb_gss3:*
- 85: gb_gss4:*
- 86: em_gss1:*
- 87: em_gss2:*
- 88: em_gss3:*
- 89: em_gss4:*
- 90: gb_gss5:*
- 91: gb_gss6:*
- 92: gb_gss7:*
- 93: gb_gss8:*
- 94: gb_gss9:*
- 95: em_gss5:*
- 96: em_gss6:*
- 97: em_gss7:*
- 98: em_gss8:*
- 99: em_gss9:*
- 100: em_gss10:*
- 101: em_gss11:*
- 102: gb_gss10:*
- 103: gb_gss11:*
- 104: em_gss12:*
- 105: gb_gss12:*
- 106: gb_gss13:*
- 107: gb_gss14:*
- 108: gb_gss15:*
- 109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result % Query

No.	Score	Match	Length	DB	ID	Description
c 1	93.8	0.3	109	30	AA243009	zr25h02.s
c 2	92.2	0.3	106	37	AA703692	ag81a10.r
c 3	91.6	0.3	106	105	AQ264176	CITBI-El-
c 4	90.8	0.3	110	106	AQ386882	RPC111-13
c 5	87.4	0.3	105	105	AQ282107	RPC111-94
c 6	86	0.3	103	38	AA807640	nx08b05.s
c 7	85.4	0.3	103	103	AQ353244	RPC1-11-3
c 8	85	0.3	110	39	AA897366	am06h02.s
c 9	85	0.3	110	94	AQ003188	RPC111-ID
c 10	84.6	0.3	107	35	AA565533	nk42b11.s
c 11	84.4	0.3	109	30	AA244173	nc05h06.s
c 12	82.6	0.3	105	28	AA078003	7h12D08 C
c 13	82.8	0.3	110	30	AA244245	nc07a04.s
c 14	82.8	0.3	110	106	AQ386882	RPC111-13
c 15	82.2	0.3	103	35	AA570438	nk63g02.s
c 16	82	0.3	103	84	B48914	RPC111-4A12
c 17	82.2	0.3	103	108	AQ334922	RPC1-11-3
c 18	82	0.3	106	63	AI991750	wt48g01.x
c 19	81.8	0.3	109	94	AQ029690	RPC111-41
c 20	82	0.3	110	32	AA369482	EST80906
c 21	81.2	0.3	104	108	AQ544583	CITBI-El-
c 22	81.2	0.3	105	109	AQ637292	RPC1-11-4
c 23	81.2	0.3	106	30	AA250812	zs06a05.s
c 24	81.4	0.3	107	35	AA583252	nn41a04.s
c 25	81.4	0.3	108	84	B65160	CIT-HSP-201
c 26	80.6	0.3	103	108	AQ535244	RPC1-11-3
c 27	80.8	0.3	106	38	AA812141	ob48h02.s
c 28	80.6	0.3	110	29	AA177157	nc02g07.s
c 29	80.2	0.3	101	35	AA583697	nn58f10.s
c 30	80.2	0.3	106	30	AA250812	zs06a05.s
c 31	80.2	0.3	109	84	B17434	345K2.TVB C
c 32	80.4	0.3	110	30	AA244245	nc07a04.s
c 33	79.8	0.3	101	33	AA381369	EST94442
c 34	79.8	0.3	102	84	B48088	RPC111-4N6.
c 35	79.8	0.3	103	108	AQ584425	RPC1-11-4
c 36	80	0.3	104	105	AQ321855	RPC111-11
c 37	80	0.3	107	33	AA385808	EST99495
c 38	80.2	0.3	108	32	AA370029	EST81584
c 39	80.2	0.3	109	84	B17434	345K2.TVB C
c 40	79.6	0.3	106	20	T55212	yb43g11.sl
c 41	79.8	0.3	107	62	AI933497	wm74d02.x
c 42	79.8	0.3	109	94	AQ029690	RPC111-41
c 43	79.4	0.3	105	61	AI832832	at72g09.x
c 44	79.4	0.3	106	44	AI249096	qbt73g09.x
c 45	79.4	0.3	106	108	AQ544957	CITBI-El-

ALIGNMENTS

RESULT 1
AA243009/c
LOCUS
DEFINITION AA243009 109 bp mRNA EST 11-MAR-1998
cDNA clone IMAGE:664467 3' similar to contains Alu repetitive element; contains element LTR1 repetitive element ;, mRNA sequence.
ACCESSION AA243009.1 GI:1873869
VERSION EST.
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 109)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Willson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Dec 3, 1996 this sequence version replaced gi:1136869.

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1127 Std Error: 0.00
Seq Primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 102.
location/Qualifiers
1. .109
/organism="Homo sapiens"
/db_xref="GDB:5426481"
/db_xref="taxon:9606"
/clone="IMAGE:664467"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/tissue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: brain; Vector: pBluescript SK-; Site: 1: EcoRI; Site: 2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Uninduced, exponentially growing neuroepithelial cells (Ntera-2/cl.D1). Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'."
BASE COUNT 19 a 30 c 30 g 30 t
ORIGIN
Query Match 0.3%; Score 93.8; DB 30; Length 109;
Best Local Similarity 93.3%; Pred. No. 0.11;
Matches 98; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 26225 CACGCCGTGTAATCCAGCACATTTGGGAGGCTGAGGTGGTGAATCAGAGTCAGGAGAT 26284
|||||
Db 109 CACGCGTGTATCCAGCACATTTGGGAGGCGGATGCGGATCAGGAGTAGGAGAT 50
Qy 26285 CAAGCACCATCTGCCACATGTTGACACCCCTCTCTACTATAAA 26329
|||||
Db 49 CAAGCACCATCTGCTAACACGGTGAACCCCGTCTCTACTATAAA 5
RESULT 2
AA703692/c
LOCUS ag81a10.r1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone
DEFINITION IMAGE:1140858 5' similar to contains Alu repetitive element; , mRNA sequence.
ACCESSION AA703692
VERSION AA703692.1 GI:2713610
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 106)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Willson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1397630.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 53.

FEATURES

Location/Qualifiers
1..106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1140858"
/clone_lib="Stratagene hNT neuron (#937233)"
/dev_stage="hNT neurons"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: phuscript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Differentiated, post mitotic hNT neurons. Average insert size: 1.5 kb; Un1-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGGAG 3' -3' adaptor sequence: 5' CTCACATTTTCTTTTCTTTT 3"

19 a 29 c 29 g 29 t

BASE COUNT

ORIGIN

Query Match 0.3%; Score 92.2; DB 37; Length 106;
Best Local Similarity 92.4%; Pred. No. 0.17;
Matches 97; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 26225 CACGCTCTAATCCAGCAGCTTTGGAGGCTGAGTGGGTGAATCAGCAGGTCAGGAGAT 26284
|||||
DB 106 CACGCTCTAATCCAGCAGCTTTGGAGGCTGAGTGGGTGAATCAGCAGGTCAGGAGAT 47
|||||

QY 26285 CAGACCATCTCTGGCCACATGGTGAACCCCGTCTCTACTAAAA 26329
|||||
DB 46 CGAGACCATCTCTGGTGAACACGCGTGAACCCCGTCTCTACTAAAA 2

RESULT 3

AQ264176 106 bp DNA GSS 27-OCT-1998
LOCUS CITBI-EI-2509A2.TF CITBI-EI Homo sapiens genomic clone 2509A2,
DEFINITION genomic survey sequence.
ACCESSION AQ264176
VERSION AQ264176.1 GI:3792743
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 106)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
JOURNAL Map Building
COMMENT Other_GSSs: CITBI-EI-2509A2.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES

Location/Qualifiers
1..106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2509A2"
/clone_lib="CITBI-EI"
/sex="male"
/cell_type="sperm"

/note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"

BASE COUNT 25 a 30 c 34 g 17 t
ORIGIN

Query Match 0.3%; Score 91.6; DB 105; Length 106;
Best Local Similarity 91.5%; Pred. No. 0.2;
Matches 97; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 26210 CCGGTGTCAGTGGCTCAGCCCTGTATCCAGCAGCTTTGGAGGTGAGTGGGTGAATC 26269
|||||
DB 1 CCGGGCGCAGAGTCTCACCCCTGTATCCAGCAGCTTTGGAGAGCCGAGCGGTGATC 60
|||||

QY 26270 ACGAGTCAGGAGATCAGACCATCTCGCCCAACATGGTGAACCC 26315
|||||
DB 61 ACGAGTCAGGAGATCAGACCCGCTCTGCTAACAATGGTGAACCC 106

RESULT 4

AQ386882/c 110 bp DNA GSS 21-MAY-1999
LOCUS RPC111-134I4.TV RPCI-11 Homo sapiens genomic clone RPCI-11-134I4,
DEFINITION genomic survey sequence.
ACCESSION AQ386882
VERSION AQ386882.1 GI:4357905
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 110)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
JOURNAL Map Building
COMMENT Other_GSSs: RPC111-134I4.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbs@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: 17
Class: BAC ends.

FEATURES

Location/Qualifiers
1..110
/organism="Homo sapiens"
/db_xref="GDB:7551267"
/db_xref="taxon:9606"
/clone="RPCI-11-134I4"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC Library"

BASE COUNT

26 a 26 c 38 g 20 t

ORIGIN

Query Match 0.3%; Score 90.8; DB 106; Length 110;
Best Local Similarity 89.1%; Pred. No. 0.25;
Matches 98; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 11717 GGGTTTCAACATGTTGGCCAGGCTGGTGTGAACCTCCTCAAGTCACCTGCC 11776

```

|||||
Db 110 GGGTTTCCACCTAGTTGTCAGGCTGGTCTTGAACCTCTTGAGCTCAAGCGATCCACCTGCC 11
QY 11777 TCAGCCTCACATAGTCTCGGAGATTACAGCGTGAGCCACCATGCTGGCC 11826
Db 50 TCAGCCTCCCAAGTACCTTGGATTACAGCGTGAGCCCATGCTCTCCCGCC 1

RESULT 5
AQ282107 105 bp DNA GSS 27-APR-1999
LOCUS AQ282107.1 RP11-94B21.TJ RP11-11 Homo sapiens genomic clone RP11-11-94B21,
DEFINITION genomic survey sequence.
ACCESSION AQ282107
VERSION AQ282107.1 GI:3907976
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 105)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RP11-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
FEATURES             Location/Qualifiers
     source
       1..105
         /organism="Homo sapiens"
         /db_xref="GDB:7535756"
         /db_xref="taxon:9606"
         /clone="RP11-11-94B21"
         /clone_lib="RP11-11"
         /sex="Male"
         /cell_type="Lymphocytes"
         /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RP11 Human Male BAC Library"
BASE COUNT 26 a 31 c 30 g 18 t
ORIGIN
Query Match 0.3%; Score 87.4; DB 105; Length 105;
Best Local Similarity 89.5%; Pred. No. 0.6;
Matches 94; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 26222 GCTCAGCGCTGTATCCAGCAGCTTTGGAGAGCTGAGTGGTGAATCAGGATCAGGA 26281
|||||
Db 1 GCTCAGCGCTGTATCCAGCAGCTCTGGAGGCCAAGGTGGGTGGATCAGGAGGCATGA 60
QY 26282 GATCAAGACCATCCCGCCACATCGTGAACCCCGTCTCTACTA 26326
|||||
Db 61 GTACGAGACCAAGCCCTGACCAACATGGTGAACCCCGTCTCTACTA 105

RESULT 6
AA807640/c 103 bp mRNA EST 05-MAR-1998
LOCUS AA807640.1 RP11-11-317H22.TV RP11-11 Homo sapiens genomic clone
DEFINITION RP11-11-317H22, genomic survey sequence.
ACCESSION AQ535244
VERSION AQ535244.1 GI:4846934
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

ACCESSION AA807640
VERSION AA807640.1 GI:2877108
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2151346.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 774 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 87.
FEATURES             Location/Qualifiers
     source
       1..103
         /organism="Homo sapiens"
         /db_xref="taxon:9606"
         /clone="IMAGE:1255473"
         /clone_lib="NCI-CGAP_GC3"
         /tissue_type="pooled germ cell tumors"
         /lab_host="DH10B"
         /note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is not normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 19 a 27 c 30 g 27 t
ORIGIN
Query Match 0.3%; Score 86; DB 38; Length 103;
Best Local Similarity 90.2%; Pred. No. 0.87;
Matches 92; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 26224 TCAGCGCTGTATCCAGCAGCTTTGGGAGGTGAGTGGGTGAATCAGGATCAGGAGA 26283
|||||
Db 103 TCACACCTGTATCCAGCAGCTTTGGGAGGCGGAGTGGAGGATCACAAGGTTCAGGAGA 44
QY 26284 TCAGACCATCTCTGGCCACATCGTGAACCCCGTCTCTACT 26325
|||||
Db 43 TCGAGACCATCTCTGGCTAACACGGTGAACCCCATCTCTACT 2

RESULT 7
AQ535244 103 bp DNA GSS 18-MAY-1999
LOCUS AQ535244.1 RP11-11-317H22.TV RP11-11 Homo sapiens genomic clone
DEFINITION RP11-11-317H22, genomic survey sequence.
ACCESSION AQ535244
VERSION AQ535244.1 GI:4846934
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

```



```

source      1. .110
            /organism="Homo sapiens"
            /db_xref="GDB:7500081"
            /db_xref="taxon:9606"
            /clone="RPC1-11-1D10"
            /clone_lib="RPC1-11"
            /sex="Male"
            /cell_type="Lymphocytes"
            /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
            RPC111 Human Male BAC Library"
BASE COUNT  22 a 27 c 26 g 35 t
ORIGIN

Query Match      0.3%; Score 85; DB 94; Length 110;
Best Local Similarity 86.2%; Pred. No. 1.1;
Matches 94; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 26230 CTGTAATCCAGCACTTTGGGAGGCTGAGGTGGTGAATCACGAGGTCAAGAGATCAAGA 26289
Db 110 CTGTAATCCAGCACTTTGGGAGGCTGCGGCAGGTGGATCATGAGTCAAGAGATCGGGA 51
Qy 26290 CCATCTCGGCCAACATGTTGGAACCCCGTCTCTACTAAATACAAAAA 26338
Db 50 CCATCTCGGCCAACATGTTGTAACCCCTGTATCTACTAAAAATACAAAAA 2

RESULT 10
AA565533/c
LOCUS      AA565533      107 bp      mRNA      EST      08-SEP-1997
DEFINITION nk42b11.s1 NCI-CGAP_GC2 Homo sapiens cDNA clone IMAGE:1016157 3'
            similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION  AA565533
VERSION     AA565533.1 GI:2337172
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 107)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT    On Sep 12, 1996 this sequence version replaced gi:1393355.
            Contact: Robert Strausberg, Ph.D.
            Tel.: (301) 496-1550
            Email: Robert.Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            CDNA Library Preparation: Stratagene, Inc., David B. Krizman,
            Ph.D.
            CDNA Library Arraying: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 1661 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 87.
Location/Qualifiers
1. .107
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1016157"
/clone_lib="NCI-CGAP_GC2"
/tissue_type="germ cell tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: Bluescript SK-; Site.1: EcoRI; Site.2:
XhoI; Cloned unidirectionally. Primer: Oligo dt. Bulk
germ cell tumor. 5' adaptor sequence: 5' GAATTCGGCAGCAG 3'
3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTTTT 3'

FEATURES
source
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 90.
Location/Qualifiers
1. .109
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1007291"
/clone_lib="NCI-CGAP_Pr1"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected, histologically normal
prostate epithelial cells. Double-stranded cDNA was
ligated to EcoRI adaptors, 5 cycles of PCR applied to the
cDNA with an adaptor-specific primer, and the resulting
PCR product subcloned into pAMP10 by the UDG-cloning
method (Life Technologies). Average insert size is 600
bp. NOTE: Not directionally cloned. This library was
constructed by David Krizman."
BASE COUNT  28 a 28 c 31 g 22 t
ORIGIN

```

```

BASE COUNT  22 a 34 c 26 g 25 t
ORIGIN

Query Match      0.3%; Score 84.6; DB 35; Length 107;
Best Local Similarity 86.9%; Pred. No. 1.2;
Matches 93; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 15950 TGGCACATGCTTAATCCAGCTACTTTGGGAGGCTCAAGCAAGAATCAGTGAACCG 15909
Db 107 TGGTGTGCTGTAATCCAGCTACTCTAGGAGGCTGAGGCAGAGAAATCAGTGAACCT 48
Qy 15910 GGGAGGCGAGGTTGCGAGTGCAGTTCAGATTTCGCCACTACACTACAG 15956
Db 47 GGGAGGCGAGGTTGCGAGTGCAGTTCAGATTTCGCCACTACACTCCAG 1

RESULT 11
AA244173/c
LOCUS      AA244173      109 bp      mRNA      EST      20-AUG-1997
DEFINITION nc05h06.s1 NCI-CGAP_Pr1 Homo sapiens cDNA clone IMAGE:1007291
            similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION  AA244173
VERSION     AA244173.1 GI:1874876
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 109)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT    On Nov 29, 1993 this sequence version replaced gi:430513.
            Contact: Robert Strausberg, Ph.D.
            Tel.: (301) 496-1550
            Email: Robert.Strausberg@nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui,
            M.D., Michael Emmert-Buck, M.D., Ph.D.
            CDNA Library Preparation: David B. Krizman, Ph.D.
            CDNA Library Arraying by: Genome Systems Inc., Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 90.
Location/Qualifiers
1. .109
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1007291"
/clone_lib="NCI-CGAP_Pr1"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected, histologically normal
prostate epithelial cells. Double-stranded cDNA was
ligated to EcoRI adaptors, 5 cycles of PCR applied to the
cDNA with an adaptor-specific primer, and the resulting
PCR product subcloned into pAMP10 by the UDG-cloning
method (Life Technologies). Average insert size is 600
bp. NOTE: Not directionally cloned. This library was
constructed by David Krizman."
BASE COUNT  28 a 28 c 31 g 22 t
ORIGIN

```

Query Match	0.3%	Score 84.4	DB 30	Length 109
Best Local Similarity	89.2%	Pred. No. 1.3		
Matches	91	Conservative	0	Mismatches 11; Indels 0; Gaps 0;
QY 11671	CCCTACCACCATGCCCGCTAAATTTTGTATTTTGTAGTAGACACAGGGTTTCACCATGT	11730		
Db	109	CACCCACCACCATGCCCTGGCTAATTTCTGTATTTTGTAGTAGACAGGGTTTCACCATGT	50	
QY 11731	TGCCAGGCTGGTGTGAACCTCCTGACCTCAAGTGATCCACC	11772		
Db	49	TGGCCAGGCTGGTGCATGAACCTCCTGACCGTAGTGTGATCCACC	8	
RESULT 12				
AA078003/c				
LOCUS	105 bp	mRNA	EST	24-SEP-1999
DEFINITION	7H12D08 Chromosome 7 HeLa cDNA Library Homo sapiens cDNA clone			
ACCESSION	AA078003			
VERSION	AA078003.1 GI:1837477			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
AUTHORS	Eutheria; Primates; Catarrhini; Hominidae; Homo.			
TITLE	1 (bases 1 to 105)			
JOURNAL	Touchman,J.W., Bouffard,G.G., Weintraub,L.A., Idol,J.R., Wang,L.,			
MEDLINE	Robbins,C.M., Nussbaum,J.C., Lovett,M. and Green,E.D.			
COMMENT	2006 expressed-sequence tags derived from human chromosome 7-enriched cDNA libraries			
	Genome Res. 7 (3), 281-292 (1997)			
	97228905			
	On Apr 14, 1993 this sequence version replaced gi:693433.			
	Contact: Eric D. Green			
	Genome Technology Branch			
	National Human Genome Research Institute/NIH			
	49 Convent Dr., MSC431, Building 49, Room 2A08, Bethesda, MD 20892			
	Tel: 3014020201			
	Fax: 3014024735			
	Email: egreen@nhgri.nih.gov			
	Plate: 12 row: D column: 08			
	Seq primer: -21M13 (AB1).			
FEATURES	Location/Qualifiers			
source	1..105			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="7H12D08"			
	/clone_lib="Chromosome 7 HeLa cDNA Library"			
	/sex="female"			
	/cell_line="HeLa cell line; ATCC"			
	/lab_host="E. coli strain DH5 alpha"			
	/note="Vector: pAMP10; cDNA was generated from cytoplasmic RNA using a mixture of random DNA hexamers and oligo(dT). From this pool of cDNA, human chromosome 7-enriched cDNA was isolated by direct cDNA selection using chromosome 7 genomic DNA (cosmids). The resulting direct-selected cDNA was cloned into a plasmid vector using a non-directional uracil DNA glycosylase (UDG)-mediated cloning strategy."			
BASE COUNT	20 a	33 c	23 g	29 t
ORIGIN				
Query Match	0.3%	Score 82.6	DB 28	Length 105
Best Local Similarity	86.7%	Pred. No. 2.1		
Matches	91	Conservative	0	Mismatches 14; Indels 0; Gaps 0;
QY 15828	TACAACAATTAGCTGGGCATTGTGGCACATGCGCTGTAATCCACGCTACTTGGGAGGCTGA	15887		
Db	105	TACAAAAATTAGCTGGGCATTGTGGCCACGCATGTAAATCCAGCTACTTGGGAGGCTGA	46	
QY 15888	AGCAAGAATATCACTTGAACCGGGGAGGACAGGTTGTCAGTGAGC	15932		

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2000, 11:34:48 ; Search time 372.06 Seconds
(without alignments)
10131.958 Million cell updates/sec

Title: US-08-852-495C-2_COPY_140000_169000
Perfect score: 29001
Sequence: 1 GCTTTTGACAAAGGTCTCAA.....TCCTTCAGGAGTTACTTCTA 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/6_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq.*
7: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	76.8	0.3	105	4	US-08-481-658B-65
C 2	76.8	0.3	105	4	US-08-477-504A-65
C 3	76.8	0.3	105	4	US-08-486-756A-65
C 4	76.8	0.3	105	4	US-08-485-862B-65
C 5	76.8	0.3	105	5	US-08-787-739-65
C 6	71.2	0.2	105	4	US-08-481-658B-65
C 7	71.2	0.2	105	4	US-08-477-504A-65
C 8	71.2	0.2	105	4	US-08-486-756A-65
C 9	71.2	0.2	105	4	US-08-485-862B-65
C 10	71.2	0.2	105	5	US-08-787-739-65
C 11	65.4	0.2	84	3	US-08-454-557C-91
C 12	65.4	0.2	84	4	US-08-340-426D-91
C 13	65.4	0.2	84	4	US-08-450-673C-91
C 14	65.4	0.2	84	6	PCT-US95-17111A-91
C 15	62.8	0.2	98	1	US-08-088-658-42
C 16	62.8	0.2	98	4	US-08-471-907A-42
C 17	60.4	0.2	78	3	US-08-454-557C-70
C 18	60.4	0.2	78	4	US-08-340-426D-70
C 19	60.4	0.2	78	4	US-08-450-673C-70
C 20	60.4	0.2	78	6	PCT-US95-17111A-70
C 21	59.6	0.2	84	2	US-08-664-596B-3
C 22	59.6	0.2	84	2	US-08-738-367-3
C 23	58.2	0.2	80	2	US-07-920-281C-25
C 24	57	0.2	76	3	US-08-454-557C-69
C 25	57	0.2	76	4	US-08-340-426D-69
C 26	57	0.2	76	4	US-08-450-673C-69
C 27	57	0.2	76	6	PCT-US95-17111A-69

28	55.8	0.2	69	2	US-08-702-344-7	Sequence 7, Appl
29	55.8	0.2	69	2	US-08-702-344-22	Sequence 92, Appl
C 30	56	0.2	85	3	US-08-454-557C-92	Sequence 22, Appl
C 31	56	0.2	85	4	US-08-340-426D-92	Sequence 92, Appl
C 32	56	0.2	85	4	US-08-450-673C-92	Sequence 92, Appl
C 33	56	0.2	85	6	PCT-US95-17111A-92	Sequence 92, Appl
C 34	55.6	0.2	78	3	US-08-454-557C-70	Sequence 70, Appl
C 35	55.6	0.2	78	4	US-08-340-426D-70	Sequence 70, Appl
C 36	55.6	0.2	78	4	US-08-450-673C-70	Sequence 70, Appl
C 37	55.6	0.2	78	6	PCT-US95-17111A-70	Sequence 70, Appl
C 38	55.2	0.2	60	3	US-08-454-557C-57	Sequence 57, Appl
C 39	55.2	0.2	60	4	US-08-340-426D-57	Sequence 57, Appl
C 40	55.2	0.2	60	4	US-08-450-673C-57	Sequence 57, Appl
C 41	55.2	0.2	60	6	PCT-US95-17111A-57	Sequence 57, Appl
C 42	55	0.2	84	3	US-08-454-557C-91	Sequence 91, Appl
C 43	55	0.2	84	4	US-08-340-426D-91	Sequence 91, Appl
C 44	55	0.2	84	4	US-08-450-673C-91	Sequence 91, Appl
C 45	55	0.2	84	6	PCT-US95-17111A-91	Sequence 91, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-65/c
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA: US/08/481,658B
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.3%; Score 76.8; DB 4; Length 105;
Best Local Similarity 83.7%; Pred. No. 1.1e-06;

Matches 87; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 26235 ATCCAGCAGCTTTGGAGGCTGGGTGAATCAGAGGTGAGGATCAAGACCATC 26294
|||||
Db 105 ATCCAGCAGCTTTGGAGGCGGAGGCTGGGTGATCAAGGTGAGGAGTTTGAGAGCAGC 46
|||||
QY 26295 CTGCGCAACATGTTGAACCCCGCTCTCTACTAAATAACAAAAA 26338
|||||
Db 45 CTGCGCAATATGTTGAACCCCTGCTCTACTAAAGATGTAATAA 2

RESULT 2

US-08-477-504A-65/c
; Sequence 65, Application US/08477504A
; Patent No. 5972353
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,504A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-477-504A-65

Query Match 0.3%; Score 76.8; DB 4; Length 105;
Best Local Similarity 83.7%; Pred. No. 1.le-06;
Matches 87; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 26235 ATCCAGCAGCTTTGGAGGCTGGGTGAATCAGAGGTGAGGATCAAGACCATC 26294
|||||
Db 105 ATCCAGCAGCTTTGGAGGCGGAGGCTGGGTGATCAAGGTGAGGAGTTTGAGAGCAGC 46
|||||
QY 26295 CTGCGCAACATGTTGAACCCCGCTCTCTACTAAATAACAAAAA 26338
|||||
Db 45 CTGCGCAATATGTTGAACCCCTGCTCTACTAAAGATGTAATAA 2

RESULT 3

US-08-486-756A-65/c
; Sequence 65, Application US/08486756A
; Patent No. 5981711
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,756A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-486-756A-65

Query Match 0.3%; Score 76.8; DB 4; Length 105;
Best Local Similarity 83.7%; Pred. No. 1.le-06;
Matches 87; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 26235 ATCCAGCAGCTTTGGAGGCTGGGTGAATCAGAGGTGAGGATCAAGACCATC 26294
|||||
Db 105 ATCCAGCAGCTTTGGAGGCGGAGGCTGGGTGATCAAGGTGAGGAGTTTGAGAGCAGC 46
|||||
QY 26295 CTGCGCAACATGTTGAACCCCGCTCTCTACTAAATAACAAAAA 26338
|||||
Db 45 CTGCGCAATATGTTGAACCCCTGCTCTACTAAAGATGTAATAA 2

RESULT 4

US-08-485-862B-65/c
; Sequence 65, Application US/08485862B
; Patent No. 5989838
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court

```
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,862B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-0727
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-485-862B-65

Query Match          0.3%; Score 76.8; DB 4; Length 105;
Best Local Similarity 83.7%; Pred. No. 1.1e-06;
Matches 87; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 26235 ATCCGAGCAGCTTTGGGAGCGCTGGTGAATCAGAGGTGAGGATCAAGACCATC 26294
Db 105 ATCCGAGCAGCTTTGGGAGCGCGAGCTGGTGAATCAGAGGTGAGGATCAAGACCATC 46

Qy 26295 CTGGCCAAATGTTGGTGAACCCGCTCTCTACTAAATACAAAAA 26338
Db 45 CTGGCCAAATGTTGGTGAACCCGCTCTCTACTAAATGTTAAAAA 2

RESULT 5
US-08-787-739-65/c
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-787-739-65

Query Match          0.3%; Score 76.8; DB 5; Length 105;
Best Local Similarity 83.7%; Pred. No. 1.1e-06;
Matches 87; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 26235 ATCCGAGCAGCTTTGGGAGCGCTGGTGAATCAGAGGTGAGGATCAAGACCATC 26294
Db 105 ATCCGAGCAGCTTTGGGAGCGCGAGCTGGTGAATCAGAGGTGAGGATCAAGACCATC 46

Qy 26295 CTGGCCAAATGTTGGTGAACCCGCTCTCTACTAAATACAAAAA 26338
Db 45 CTGGCCAAATGTTGGTGAACCCGCTCTCTACTAAATGTTAAAAA 2

RESULT 6
US-08-481-658B-65
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.2%; Score 71.2; DB 4; Length 105;
Best Local Similarity 85.8%; Pred. No. 1.4e-05;
Matches 91; Conservative 0; Mismatches 13; Indels 2; Gaps 1;
QY 18903 TTTTGTATTTTAGTAGAGATAGGTTTCACAAATGCTGCCAGGCTGCTCAAACTCC 18962
DB 2 TTTTACATCTTTAGTAGAGACAGGTTTCCACCATATTTGCCAGGCTGCTCAAACTCC 61
QY 18963 TGCCTCAAGTGATCCCTCGCCCGCCCTCCCAATGCTGCTGGAT 19008
DB 62 TGACCT--TGTGATCCACGACCTCGGCCCTCCCAAGTGTCTGGAT 105

RESULT 7
US-08-477-504A-65
Sequence 65, Application US/08477504A
Patent No. 5972353
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-477-504A-65
Query Match 0.2%; Score 71.2; DB 4; Length 105;
Best Local Similarity 85.8%; Pred. No. 1.4e-05;
Matches 91; Conservative 0; Mismatches 13; Indels 2; Gaps 1;
QY 18903 TTTTGTATTTTAGTAGAGATAGGTTTCACAAATGCTGCCAGGCTGCTCAAACTCC 18962
DB 2 TTTTACATCTTTAGTAGAGACAGGTTTCCACCATATTTGCCAGGCTGCTCAAACTCC 61
QY 18963 TGCCTCAAGTGATCCCTCGCCCGCCCTCCCAATGCTGCTGGAT 19008
DB 62 TGACCT--TGTGATCCACGACCTCGGCCCTCCCAAGTGTCTGGAT 105
RESULT 8
US-08-486-756A-65
Sequence 65, Application US/08486756A
Patent No. 5981711
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,756A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-486-756A-65

Query Match 0.2%; Score 71.2; DB 4; Length 105;
Best Local Similarity 85.8%; Pred. No. 1.4e-05;
Matches 91; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 18903 TTTTGTATTTTAGTAGAGATAGGTTTTCACAAATGCTGGCCAGGCTGCTCTCAAACTCC 18962
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 2 TTTTACATCTTTTAGTAGAGACAGGTTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

QY 18963 TGGCTCAAGTGATCTCTCGCTGGCTCCCAATGTGCTGGGAT 19008
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 62 TGACCT--TGTGATCCACGAGCTGGCTCCCAAGTGCTGGGAT 105

RESULT 9
US-08-485-862B-65
Sequence 65, Application US/08485862B
Patent No. 5989838
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485.862B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-485-862B-65

Query Match 0.2%; Score 71.2; DB 4; Length 105;
Best Local Similarity 85.8%; Pred. No. 1.4e-05;
Matches 91; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 18903 TTTTGTATTTTAGTAGAGATAGGTTTTCACAAATGCTGGCCAGGCTGCTCTCAAACTCC 18962
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 2 TTTTACATCTTTTAGTAGAGACAGGTTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61
QY 18963 TGGCTCAAGTGATCTCTCGCTGGCTCCCAATGTGCTGGGAT 19008
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 62 TGACCT--TGTGATCCACGAGCTGGCTCCCAAGTGCTGGGAT 105

RESULT 10
US-08-787-739-65
Sequence 65, Application US/08787739
Patent No. 6027887
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 96
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 369 Pine Street, Suite 610
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/787,739
FILING DATE: 24-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/486,756
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/481,658
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,862
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,863
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,077
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.4
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-787-739-65

```
Query Match          0.2%; Score 71.2; DB 5; Length 105;
Best Local Similarity 85.8%; Pred. No. 1.4e-05;
Matches 91; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

Qy 18903 TTTTGTATTTTGTAGATAGAGTTTCAATGCTGGCCAGGCTGGTCTCAAACTCC 18962
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 TTTTACATCTTTAGTAGAGACAGGTTTCAACATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy 18963 TCCCTCAAGTCATCTCTGCTGGCTCCCAATGCTGCTGGGAT 19008
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 TGACCT--TGTGATCCACCAGCTGGCTCCCAAGTCTGCTGGAT 105

RESULT 11
US-08-454-557C-91/c
; Sequence 91, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.38400003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-454-557C-91

Query Match          0.2%; Score 65.4; DB 3; Length 84;
Best Local Similarity 86.7%; Pred. No. 0.00017;
Matches 72; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 26225 CACGCGCTGTATCCAGCACATTTGGAGGCTGAGTGGTGAATCAGGAGTCAAGAGAT 26284
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 83 CACGCTTGTATCCAGCACATTTGGAGGCTGAGTGGTGAATCAGGAGTCAAGAGAT 24

Qy 26285 CAAGACCATCTCGGCCAACATGG 26307
      ||||| ||||| ||||| |||||
Db .23 CGACACCAGCCTGTGATGAACATGG 1

RESULT 12
US-08-340-426D-91/c
; Sequence 91, Application US/08340426D
; Patent No. 5948634
```

```
GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.38400002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-340-426D-91

Query Match          0.2%; Score 65.4; DB 4; Length 84;
Best Local Similarity 86.7%; Pred. No. 0.00017;
Matches 72; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 26225 CACGCGCTGTATCCAGCACATTTGGAGGCTGAGTGGTGAATCAGGAGTCAAGAGAT 26284
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 83 CACGCTTGTATCCAGCACATTTGGAGGCTGAGTGGTGAATCAGGAGTCAAGAGAT 24

Qy 26285 CAAGACCATCTCGGCCAACATGG 26307
      ||||| ||||| ||||| |||||
Db 23 CGACACCAGCCTGTGATGAACATGG 1

RESULT 13
US-08-450-673C-91/c
; Sequence 91, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```


Search completed: June 20, 2000, 02:28:19
Job time: 480829 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2000, 02:22:00 ; Search time 29135.9 seconds
(without alignments)
-968.286 Million cell updates/sec

Title: US-08-852-495C-2_COPY_168000_197000
Perfect score: 29001
Sequence: 1 ATTTACAGATGGAGAAACCA.....GGATTAGGATCATGATCTC 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*

- 1: gb_bal:*
- 2: gb_ba2:*
- 3: gb_om:*
- 4: gb_ov:*
- 5: gb_pat:*
- 6: gb_ph:*
- 7: gb_pl1:*
- 8: gb_pl2:*
- 9: gb_pr1:*
- 10: gb_pr2:*
- 11: gb_pr3:*
- 12: gb_ro:*
- 13: gb_sts:*
- 14: gb_sy:*
- 15: gb_un:*
- 16: gb_v1:*
- 17: em_fun:*
- 18: em_hum1:*
- 19: em_hum2:*
- 20: em_in:*
- 21: em_om:*
- 22: em_or:*
- 23: em_ov:*
- 24: em_pat:*
- 25: em_ph:*
- 26: em_pl:*
- 27: em_ro:*
- 28: em_sts:*
- 29: em_sy:*
- 30: em_un:*
- 31: em_v1:*
- 32: gb_htg1:*
- 33: gb_htg2:*
- 34: gb_in1:*
- 35: gb_in2:*
- 36: em_ba1:*
- 37: em_ba2:*
- 38: em_hum3:*
- 39: em_hum4:*
- 40: gb_pr4:*
- 41: gb_htg3:*
- 42: gb_htg4:*
- 43: gb_htg5:*
- 44: gb_htg6:*

- 45: gb_htg7:*
- 46: em_htg1:*
- 47: em_htg2:*
- 48: em_htg3:*
- 49: em_hum5:*
- 50: gb_pl3:*
- 51: gb_pr5:*
- 52: gb_htg8:*
- 53: gb_htg9:*
- 54: gb_htg10:*
- 55: gb_htg11:*
- 56: gb_htg12:*
- 57: gb_htg13:*
- 58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
c 1	96.8	0.3	108	10	HSIDLRLN2	X05250 Human LDL-r
c 2	92	0.3	108	10	HSIDLRLN2	X05250 Human LDL-r
c 3	86.2	0.3	108	10	HSIDLRLD1	X05249 Human LDL-r
c 4	86.2	0.3	108	10	HSIDLRLD2	X05251 Human LDL-r
c 5	83.4	0.3	107	9	HUMALCE162	M87924 Human carci
c 6	83	0.3	108	10	HSIDLRLD1	X05249 Human LDL-r
c 7	83	0.3	108	10	HSIDLRLD2	X05251 Human LDL-r
c 8	82.2	0.3	103	9	HUMALCE221	M87896 Human carci
c 9	81	0.3	108	11	HSU67803	U67803 Human small
c 10	81	0.3	108	11	HSU67808	U67808 Human small
c 11	78	0.3	107	9	HUMALCE162	M87924 Human carci
c 12	77.6	0.3	103	9	HUMALCE221	M87896 Human carci
c 13	76.6	0.3	108	10	HSIDLRLI12	X05248 Human LDL-r
c 14	76	0.3	103	13	HS8IC8R	X57789 Human sequ
c 15	75	0.3	110	11	HSU67807	U67807 Human small
c 16	74.4	0.3	104	9	HUMALCE272	M87899 Human carci
c 17	74.6	0.3	108	11	HSU67804	U67804 Human small
c 18	73.4	0.3	91	13	HUMUT8164A	L30244 Human STS U
c 19	73.6	0.3	97	9	HUMDLRLA2	M14180 Human low d
c 20	73.4	0.3	110	11	HSU67807	U67807 Human small
c 21	72.8	0.3	103	13	HS8IC8R	X57789 Human sequ
c 22	73	0.3	108	13	G43535	G43535 WIAF-2393-S
c 23	72.4	0.2	91	13	HUMUT8164A	L30244 Human STS U
c 24	72.4	0.2	110	9	HUMALCE43	M87900 Human carci
c 25	72	0.2	97	9	HUMDLRLA2	M14180 Human low d
c 26	72.2	0.2	108	9	HUMDLID03M5	D16965 Human HepG2
c 27	72	0.2	108	10	HSIDLRLI12	X05248 Human LDL-r
c 28	71.4	0.2	107	11	HSU67806	U67806 Human small
c 29	70.8	0.2	90	9	HUMDLRLFL	K03555 Human low d
c 30	70.8	0.2	100	13	HUMUT931A	L31299 Human STS U
c 31	69.4	0.2	108	13	G43535	G43535 WIAF-2393-S
c 32	68.6	0.2	84	5	AR051521	AR051521 Sequence
c 33	68.8	0.2	106	13	G32743	G32743 A009P31 Hum
c 34	68.4	0.2	95	13	HUMUT8002B	L30176 Human STS U
c 35	68.6	0.2	97	9	HUMDLRLD1	M14179 Human famli
c 36	67.8	0.2	108	9	HUMDLID03M5	D16965 Human HepG2
c 37	67.4	0.2	79	10	S73203	S73203 ALL-1 (tand
c 38	67.6	0.2	99	13	HUMUT7692A	L30306 Human STS U
c 39	67.8	0.2	108	3	AF185109S1	AF185109 Lasiorhin
c 40	67.4	0.2	102	13	G32906	G32906 A009W09 Hum
c 41	67	0.2	97	9	HUMDLRLD1	M14179 Human famli
c 42	65.8	0.2	97	9	HUMDLRLA1	M14178 Human low d
c 43	65.2	0.2	95	10	HSSTHPKIB	X66361 H.sapiens m
c 44	65.4	0.2	100	13	HUMUT931A	L31299 Human STS U
c 45	65.6	0.2	106	13	G32743	G32743 A009P31 Hum

ALIGNMENTS


```

SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 108)
AUTHORS     Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
TITLE       Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
JOURNAL     Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE     87161901
COMMENT     *source: hypercholesterol aemia
            See X05250 for corresponding normal gene sequence
            In the defective LDL-receptor gene the deletion occurred between two
            alu-repetitive sequences, that are in the same direction, the
            deletion eliminates exons 13 and 14 and changes the reading frame
            of the resulting spliced mRNA.
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES    Location/Qualifiers
            source          1..108
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /cell_type="blood leukocytes from a patient with familial"
                        /note="intron XIV fragment"
            intron          1..108
            BASE COUNT     28 a 20 c 40 g 20 t
            ORIGIN
            Query Match      0.3%; Score 86.2; DB 10; Length 108;
            Best Local Similarity 87.9%; Pred. No. 4e-05;
            Matches 94; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 24902 TTGGCTCACTGCAACTCGGCTCTCTGGTTCAAGCAATTCACGCTCCGCTCCCGA 24961
DB 107 TCGCCTCACCACAACCTGCTGCTCTGGTTCAAAACCAATTTTCCTGCCCTCCCGA 48
QY 24962 GTAGCTGGGATTACAGGCACATGCCACCATGACTGGCTAATTTTGT 25008
DB 47 GTAGCTGGGATTACAGGCACCTGCCACACGCTGGCTAATTTTGT 1
RESULT 5
LOCUS      HUMALCE162 107 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE162.
ACCESSION M87924
VERSION M87924.1 GI:174871
KEYWORDS   Alu repeat.
SOURCE     Homo sapiens male embryo carcinoma CDNA to other RNA.
ORGANISM   Homo sapiens
REFERENCE   1 (bases 1 to 107)
AUTHORS     Sinnett,D., Richer,C., Daxagon,J.-M. and Labuda,D.
TITLE       Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
JOURNAL     J. Mol. Biol. (1992) in press
FEATURES    Location/Qualifiers
            source          1..107
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /cell_line="NTera2D1"
                        /dev_stage="embryo"
                        /sex="male"
                        /tissue_type="carcinoma"
            BASE COUNT     28 a 30 c 35 g 14 t
            ORIGIN
            Query Match      0.3%; Score 83.4; DB 9; Length 107;
            Best Local Similarity 89.1%; Pred. No. 0.00012;

```

```

Matches 90; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 19131 GCACAGAGTAACCTTGAACCCAGGAGGACAGATTCAGCTGAGCTGCGCCACT 19190
DB 5 GCACAGAGTAATGGCTGAACCCGGAGGCGGAGCTTGCAGTGAGCGAGATCGCGCCACT 64
QY 19191 GCATTCCAGCCCTGGGACAGACGAGCGAGACTCCATCTCAAAA 19231
DB 65 GCATCCAGCGCTGGGACAGACGAGCGAGACTCCGCTCAAAA 105
RESULT 6
LOCUS      HSLDLRD1 108 bp DNA PRI 20-MAY-1992
DEFINITION Human LDL-receptor mutated gene with intron 12 deletion junction.
ACCESSION X05249
VERSION X05249.1 GI:34335
KEYWORDS   Alu repetitive sequence; low density lipoprotein receptor.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Homnidae; Homo.
            REFERENCE      1 (bases 1 to 108)
            AUTHORS        Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
                        Williamson,R. and Humphries,S.
            TITLE          Unequal crossing-over between two alu-repetitive DNA sequences in
                        the low-density-lipoprotein-receptor gene. A possible mechanism for
                        the defect in a patient with familial hypercholesterolaemia
            JOURNAL        Eur. J. Biochem. 164 (1), 77-81 (1987)
            MEDLINE        87161901
            COMMENT        *source: hypercholesterol aemia
                        See X05248 for corresponding normal gene sequence
                        In the defective LDL-receptor gene the deletion occurred between two
                        alu-repetitive sequences, that are in the same direction, the
                        deletion eliminates exons 13 and 14 and changes the reading frame
                        of the resulting spliced mRNA.
                        Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES    Location/Qualifiers
            source          1..108
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /cell_type="blood leukocytes from a patient with familial"
                        /note="deletion junction region intron 12/ intron 15"
            BASE COUNT     20 a 40 c 20 g 28 t
            ORIGIN
            Query Match      0.3%; Score 83; DB 10; Length 108;
            Best Local Similarity 86.0%; Pred. No. 0.00014;
            Matches 92; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 19072 ACAAAATTTAGCCAGCGCTGGTGGCATCTGCTGTAGTCCAGCTACTCGGGACTGAG 19131
DB 108 ACAAAATTTAGCCAGCGCTGGTGGCATCTGCTGTAGTCCAGCTACTCGGGAGGCTGAG 49
QY 19132 GCAGGAGATCACTTGAACCCAGGAGCGAGATTCAGTGAGCTCA 19178
DB 48 GCAGGAGAAATGTTTGAACCCAGGAGCGAGGTTGTGTGAGGCCGA 2
RESULT 7
LOCUS      HSLDLRD2 108 bp DNA PRI 20-MAY-1992
DEFINITION Human LDL-receptor mutated gene with intron 14 deletion junction.
ACCESSION X05251
VERSION X05251.1 GI:34336
KEYWORDS   Alu repetitive sequence; low density lipoprotein receptor.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Homnidae; Homo.
            REFERENCE      1 (bases 1 to 108)

```

AUTHORS Horsthemke, B., Beisiegel, U., Dunning, A., HAVINGA, J. R.,
 Williamson, R., and Humphries, S.

TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
 the low-density-lipoprotein-receptor gene. A possible mechanism for
 the defect in a patient with familial hypercholesterolaemia

JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901

COMMENT *source: hypercholesterol aemia
 See X05250 for corresponding normal gene sequence
 In the defective LDL-receptor gene the deletion occurred between two
 alu-repetitive sequences, that are in the same direction, the
 deletion eliminates exons 13 and 14 and changes the reading frame
 of the resulting spliced mRNA.
 Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.

FEATURES
 source
 1. .108
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_type="blood leukocytes from a patient with familial"
 1. .108
 /note="intron XIV fragment"
 28 a 20 c 40 g 20 t

BASE COUNT
ORIGIN

Query Match 0.3%; Score 83; DB 10; Length 108;
 Best Local Similarity 86.0%; Pred. No. 0.00014;
 Matches 92; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 19072 ACAAATAATAGCAGGCTGTGGCAGTCGCTAGTCCAGTACTCGGGACACTGAG 19131
 Db 1 ACAAATAATAGCAGGCTGTGGCAGTCGCTAGTCCAGTACTCGGGAGGCTGAG 60

Qy 19132 GCAGGAGATCACTGACCCAGGAGCAGAGATTGCGAGTCTGA 19178
 Db 61 GCAGGAAATGGTTGACCCAGGAGGAGGTTGTTGGTGAGCGGA 107

RESULT 8
HUMALCE221 103 bp ss-RNA PRI 15-APR-1994
LOCUS Human carcinoma cell-derived Alu RNA transcript, clone CE221.
DEFINITION M87896
ACCESSION M87896.1 GI:174874
VERSION Alu repeat.
KEYWORDS Homo sapiens male embryo carcinoma cDNA to other RNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 103)
AUTHORS Sinnott, D., Richer, C., Deragon, J.-M. and Labuda, D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of
 post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES Location/Qualifiers
 source
 1. .103
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_line="Ntera2D1"
 /dev_stage="embryo"
 /sex="male"
 /tissue_type="carcinoma"
 25 a 27 c 33 g 18 t

BASE COUNT
ORIGIN

Query Match 0.3%; Score 82.2; DB 9; Length 103;
 Best Local Similarity 87.4%; Pred. No. 0.00019;
 Matches 90; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 19097 ATCTGCGCTAGTCCAGCTACTCGGGACACTGAGGCGAGGAGTAATCACTTCAACCCAGGA 19156
 Db 1 ATCTGCGCTAGTCCAGCTACTCGGGAGGAGTAATCACTTCAACCCCGGGA 60

Qy 19157 GGCAGAGATTGCAGTGCAGCTGAGATCGCGCACATTCGATTCAG 19199
 Db 61 GCGGAGGTTGCATGAGCCGAGATCGTGCATTCGACTCCAG 103

RESULT 9
HSU67803 108 bp RNA PRI 01-AUG-1997
LOCUS Human small cytoplasmic Alu transcript.
DEFINITION U67803
ACCESSION U67803
VERSION U67803.1 GI:2289917
KEYWORDS Alu.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 108)
AUTHORS Shaikh, T.H., Roy, A.M., Kim, J., Batzer, M.A. and Deininger, P.L.
TITLE cDNAs derived from primary and small cytoplasmic Alu (scAlu)
 transcripts
JOURNAL J. Mol. Biol. 271 (2), 222-234 (1997)
MEDLINE 97415756
REFERENCE 2 (bases 1 to 108)
AUTHORS Shaikh, T.H., Kim, J., Batzer, M.A. and Deininger, P.L.
TITLE Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
 Children's Hospital of Philadelphia, 1004F Abramson Research
 Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
JOURNAL Location/Qualifiers
FEATURES source
 1. .108
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="tscAlu2"
 repeat_region 1. .108
 /note="scAlu"
 /rpt_family="Alu"
 /rpt_type="dispersed"
 23 a 39 c 30 g 16 t

BASE COUNT
ORIGIN

Query Match 0.3%; Score 81; DB 11; Length 108;
 Best Local Similarity 89.7%; Pred. No. 0.0003;
 Matches 87; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 18966 GCCTGTATCCAGCACTTTGGGAGGCCAAGCGGACAGATCAGAGTCAGGAGTTGA 19025
 Db 1 GCCTGTATCCAGCACTTTGGGAGGCCGAGCGGGGATCAGAGTCAGGAGATCGA 60

Qy 19026 GACCAGCTGACCAACATGGTGAACCCCTGTCTCTAC 19062
 Db 61 GACCATCTGCTACCAAGGTGAACCCCGTCTCTAC 97

RESULT 10
HSU67808 108 bp RNA PRI 01-AUG-1997
LOCUS Human small cytoplasmic Alu transcript.
DEFINITION U67808
ACCESSION U67808
VERSION U67808.1 GI:2289922
KEYWORDS Alu.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 108)
AUTHORS Shaikh, T.H., Roy, A.M., Kim, J., Batzer, M.A. and Deininger, P.L.
TITLE cDNAs derived from primary and small cytoplasmic Alu (scAlu)
 transcripts
JOURNAL J. Mol. Biol. 271 (2), 222-234 (1997)
MEDLINE 97415756
REFERENCE 2 (bases 1 to 108)

AUTHORS Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The Children's Hospital of Philadelphia, 1004F Abramson Research Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

FEATURES
source Location/Qualifiers
1. .108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Tscalu7"
repeat_region 1. .108
/note="scAlu"
/rpt_family="Alu"
/rpt_type="dispersed"
BASE COUNT 22 a 37 c 28 g 21 t
ORIGIN

Query Match 0.3%; Score 81; DB 11; Length 108;
Best Local Similarity 89.7%; Pred. No. 0.0003;
Matches 87; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 19866 GCCTGTAATCCAGCACTTTGGAGGCCAAGCGGCAGATCAGGAGTCAGGAGTTGA 19025
|||||
Db 1 GCCTGTAATCCAGCACTTTGGAGGCCAAGTCGGGTGGATCACAAAGTCAGGAGTTGA 60
|||||

QY 19026 GACCAGCCTGACCAACATGCTGAACCCCTCTCTAC 19062
|||||
Db 61 GACCAGCCTGACCAACATGCTGAACCTCTCTCTTC 97
|||||

RESULT 11
HUMALCE162/c
LOCUS HUMALCE162 107 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE162.
ACCESSION M87924
VERSION M87924.1 GI:174871
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 107)
AUTHORS Sinnott,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) in press
FEATURES Location/Qualifiers
source 1. .107
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Ntera2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT 28 a 30 c 35 g 14 t
ORIGIN

Query Match 0.3%; Score 78; DB 9; Length 107;
Best Local Similarity 85.3%; Pred. No. 0.00094;
Matches 87; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 20785 TTTTTCAGACAGAGTCTGACTGTGTCGCCAGGCTGAGTGGCACCACCTAGGCT 20844
|||||
Db 106 TTTTTCAGAGGAGTCTGCTGTGCGCCAGGCTGAGTGGCAGTCTCGGCT 47
|||||

QY 20845 CACTGAGCCTCTACCTCTCGGCTTCAAGCGATTCTCTGCC 20886
|||||
Db 46 CACTGCAAGCTCCGCTCCCGGTTTCACGCCATTCTCTGCC 5
|||||

RESULT 12

HUMALCE221/c
LOCUS HUMALCE221 103 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Sinnott,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) in press
FEATURES Location/Qualifiers
source 1. .103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Ntera2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT 25 a 27 c 33 g 18 t
ORIGIN

Query Match 0.3%; Score 77.6; DB 9; Length 103;
Best Local Similarity 86.0%; Pred. No. 0.0011;
Matches 86; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 12799 CTGGAGTGCAGTGCAGCATCTCAGCTCACTGAACCTCAATTCCTCAGTTCAGCGAT 12858
|||||
Db 103 CTGGAGTGCAGTGCAGCATCTCAGCTCACTGAACCTCCGCCCTCCCGGGTTCAGCGAT 44
|||||

QY 12859 TCTCGTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCA 12898
|||||
Db 43 TCTCGTGCCTTAGCTCCCGTGTAGCTGGGATTACAGGCA 4
|||||

RESULT 13
HSLDLI12
LOCUS HSLDLI12 108 bp DNA PRI 20-MAY-1992
DEFINITION Human LDL-receptor gene intron 12 fragment (normal gene) LDL = low density lipoprotein.
ACCESSION X05248
VERSION X05248.1 GI:34334
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor; repetitive sequence.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R., Williamson,R. and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901
COMMENT see X05249 for deletion junction
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES Location/Qualifiers
source 1. .108
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature complement(<1..65)
/note="Alu repeat"
intron 1. .108
/note="intron XII fragment"
BASE COUNT 21 a 38 c 20 g 29 t
ORIGIN

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 110)
Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
cDNAs derived from primary and small cytoplasmic Alu (scAlu)
transcripts
J. Mol. Biol. 271 (2), 222-234 (1997)
97415756
2 (bases 1 to 110)
Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
Direct Submission
Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
Children's Hospital of Philadelphia, 1004F Abranson Research
Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES
source
1..110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TscAlu6"
/region="repeat_region"
1..110
/note="scAlu"
/rpt_family="Alu"
/rpt_type="dispersed"
26 a 39 c 24 g 21 t
BASE COUNT
ORIGIN

Query Match 0.3%; Score 75; DB 11; Length 110;
Best Local Similarity 84.8%; Pred. No. 0.003;
Matches 84; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 17338 GTAGAGTGGGTTCGCCATGTGGCCAGCTGGTCTCGAACTCCGTGGCGCTCAAGCGAT 17397
Db 99 GGAAGATGGGGTTTCACCAAGTTTGACAGCGCTGGTCTTGAATCTCGGGCTCAAGTGAT 40

QY 17398 CCACATTCCCTTGGCGCTCCCAAAGTGCCTAAGATTACAGCG 17436
Db 39 CCACCCACTTTGGCGCTCTCAAAGTGGCTGGGATTACAGCG 1

Search completed: June 20, 2000, 17:56:59
Job time: 537503 sec

```

Search completed: June 20, 2000, 17:56:59
Job time: 537503 sec

[illegible]

```
1. .103
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
```

[illegible]

RESULT	15
LOCUS	HS067807/c
DEFINITION	Human small cytoplasmic Alu transcript.
ACCESSION	U67807
VERSION	U67807.1 GI:2289921
KEYWORDS	Alu.
SOURCE	human.
ORGANISM	Homo sapiens
PRI	01-AUG-1997

;

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	73.6	0.3	108	1	X12095		Human biallelic po
2	70	0.2	108	1	X12095		Human biallelic po
C 3	67.6	0.2	108	1	T25009		Human gene signatu
4	64	0.2	92	1	T66081		(dC-dA)n. (dG-dT)n
5	62.4	0.2	100	1	T24892		Human gene signatu
C 6	62.4	0.2	100	1	T24892		Human gene signatu
7	62.6	0.2	103	1	T26213		Human gene signatu
C 8	61.6	0.2	87	1	T21566		Human gene signatu
9	61.8	0.2	103	1	T20927		Human gene signatu
C 10	61.4	0.2	108	1	T26828		Human gene signatu
C 11	61	0.2	103	1	T26213		Human gene signatu
C 12	60.6	0.2	93	1	T22572		Human gene signatu
13	59.4	0.2	91	1	T25854		Human gene signatu
14	59.4	0.2	108	1	T26828		Human gene signatu
15	59	0.2	60	1	T65762		Repeat sequence fr
16	58	0.2	87	1	T21566		Human gene signatu
17	57.6	0.2	108	1	T25009		Human gene signatu
18	57	0.2	100	1	X12087		Human biallelic po
19	57	0.2	100	1	X12085		Human biallelic po
20	57	0.2	100	1	X12086		Human biallelic po
C 21	55.8	0.2	93	1	T24259		Human gene signatu
C 22	55.8	0.2	93	1	T24259		Human gene signatu
C 23	55.4	0.2	81	1	T24093		Human gene signatu
C 24	55.4	0.2	88	1	X39744		Microsatellite ana
C 25	55.2	0.2	95	1	T23131		Human gene signatu
C 26	55	0.2	99	1	T24420		Human gene signatu
C 27	54.6	0.2	91	1	T25854		Human gene signatu
C 28	53.6	0.2	69	1	Q2016		Probe to internal
C 29	53.6	0.2	91	1	T65740		Repeat sequence fr
C 30	53.2	0.2	73	1	Q34140		Sequence of a micr
C 31	53.4	0.2	99	1	T23728		Human gene signatu
C 32	53.4	0.2	100	1	X12087		Human biallelic po
C 33	53.4	0.2	100	1	X12085		Human biallelic po
C 34	53.4	0.2	100	1	X12086		Human biallelic po


```
RESULT 10
T26828/c
ID T26828 standard; cDNA to mRNA; 108 BP.
AC T26828:
DT 14-NOV-1996 (first entry)
DE Human gene signature HUMGS09078.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 2182; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 18 A; 33 C; 23 G; 28 T;

Query Match 0.2%; Score 61.4; DB 1; Length 108;
Best Local Similarity 79.8%; Pred. No. 0.57;
Matches 71; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 18934 ATCATATACTAGGCTGGCGGTGATGGCTCAGCGCTGTATCCAGCAGCTTTGGGAGGCC 18993
DB 90 AAGAAATAAACAGCGCGGGCGTGGCTCATGCTGTAAACCCAGCAGCTATGGGAGGCC 31

QY 18994 AAGCGGACAGATCAGGAGTTCAGGAGTT 19022
DB 30 GANACGGCGGATGACGAGTTCAGGAGAT 2

RESULT 11
T26213/c
ID T26213 standard; cDNA to mRNA; 103 BP.
AC T26213:
DT 13-NOV-1996 (first entry)
DE Human gene signature HUMGS08452.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.

PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1159; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 103 BP; 33 A; 21 C; 25 G; 23 T;

Query Match 0.2%; Score 61; DB 1; Length 103;
Best Local Similarity 75.2%; Pred. No. 0.65;
Matches 76; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 27386 TTTTGTGTTTATAGACACAGCATCTTATCTGTCCACCGCTGTAGCTCAGTGGCCAA 27445
DB 102 TTTTGTGTTTATAGACAGCATGTCTTACTCTGTGGCCAGGCTGGAGTGGGTGCCA 43

QY 27446 TCATTCTCTCACTCAGCGCTCAACTCTCTGGCTCCAGTAAT 27486
DB 42 TCATAGCTCAGCTATACACCAAACTCTCGGACTCAAGTGAT 2

RESULT 12
T22572/c
ID T22572 standard; cDNA to mRNA; 93 BP.
AC T22572:
DT 01-OCT-1996 (first entry)
DE Human gene signature HUMGS04188.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1159; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 103 BP; 33 A; 21 C; 25 G; 23 T;
```

```
CC recognising different cell types.
SQ Sequence 93 BP; 22 A; 23 C; 24 G; 22 T;

Query Match 0.28; Score 60.6; DB 1; Length 93;
Best Local Similarity 77.4%; Pred. No. 0.74;
Matches 72; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 21283 TTTTACGACAGCTCTACTCTGTCACGAGCTGGAGTGCAGCGGTGTGAACATGGCTC 21342
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 93 TTTGTAACAGGGCTCTGCAATNATCAACCAAGCTGGAGTGCAGCGGTGTGAACATGGCTC 21342
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 21343 ACTGCAGCTCAACTACTGAGCTCAAGCAATC 21375
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33 ACTGCAGCTCAACTCTCTCGGCTCAAGGGATC 1
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
T25854 standard; cDNA to mRNA; 91 BP.
AC T25854;
DT 22-OCT-1996 (first entry)
DE Human gene signature HUMGS08084.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
human; cloning; mapping; non-biased library; diagnosis; detection;
cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K. Okubo K;
WPI; 95-206931/27.
DR Identifying gene signatures in 3'-directed human cDNA library - e.g.
for diagnosis of abnormal cell function, by preparing cDNA that
reflects relative abundance of corresp. mRNA in specific human
tissues
PT
PS Claim 1; Page 1944; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
double-stranded DNA) which comprises one of the 7837 "GS" sequences
given in T19001-T26837 and which is able to hybridise to part of
human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
sequences were obtained from 3'-directed cDNA libraries prepared
from various human tissues; synthesis of cDNA was initiated from the
3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
untranslated sequence is unique to a particular mRNA species, almost
all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
is constructed so as to reflect accurately the relative abundance of
different mRNAs in the particular tissue from which it was derived.
The appearance frequency of a given GS in a cDNA library can be
determined (esp. using primers and probes derived from the GS
sequences) as a means of diagnosing abnormal cell function or for
recognising different cell types.
SQ Sequence 91 BP; 18 A; 22 C; 28 G; 18 T;

Query Match 0.28; Score 59.4; DB 1; Length 91;
Best Local Similarity 77.5%; Pred. No. 1.1;
Matches 69; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 19140 ATCACTTGAACCCAGGACGAGATTGAGTGCAGATCGGCACATGCAATCCAG 19199
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 ATCACTTGAAGCCAGGACGAGNGTTCAAGTGCAGATGAGCACTTCGCGCTCCAG 61
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 19200 CTTGGAGACAGGACGATCCATCTCA 19228
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 CTTGGTGCACGCGTGAGANNCTGCTCA 90
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
T26828
```

```
ID T26828 standard; cDNA to mRNA; 108 BP.
AC T26828;
DT 14-NOV-1996 (first entry)
DE Human gene signature HUMGS09078.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
human; cloning; mapping; non-biased library; diagnosis; detection;
cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K. Okubo K;
WPI; 95-206931/27.
DR Identifying gene signatures in 3'-directed human cDNA library - e.g.
for diagnosis of abnormal cell function, by preparing cDNA that
reflects relative abundance of corresp. mRNA in specific human
tissues
PT
PS Claim 1; Page 2182; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
double-stranded DNA) which comprises one of the 7837 "GS" sequences
given in T19001-T26837 and which is able to hybridise to part of
human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
sequences were obtained from 3'-directed cDNA libraries prepared
from various human tissues; synthesis of cDNA was initiated from the
3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
untranslated sequence is unique to a particular mRNA species, almost
all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
is constructed so as to reflect accurately the relative abundance of
different mRNAs in the particular tissue from which it was derived.
The appearance frequency of a given GS in a cDNA library can be
determined (esp. using primers and probes derived from the GS
sequences) as a means of diagnosing abnormal cell function or for
recognising different cell types.
SQ Sequence 108 BP; 18 A; 33 C; 23 G; 28 T;

Query Match 0.28; Score 59.4; DB 1; Length 108;
Best Local Similarity 72.8%; Pred. No. 1.1;
Matches 75; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 2708 GATCTCTGACCTTGTGATCCACCCGCTCAGCCTCCCAAGTCCAGGATTACAGGCAT 2767
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 GATCTCTGACCTGTGATCCGCGGNTGCGCTCCCATAGTCTGGGNTTACAGGCAT 60
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 2768 GAGCCACGCTGCCAGCCTCTTTTCTTTCTTTATAGACAAAG 2810
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 GAGCCACACACCGCGGCTGTTTATTCTTTATAACTGTACAGG 103
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
T65762 standard; DNA; 60 BP.
AC T65762;
DT 17-JUN-1997 (first entry)
DE Repeat sequence from polymorphic marker clone Mfd67.
KW Polymorphism; repeat sequence; genetic marker; primer; amplification;
PCR; polymerase chain reaction; paternity; maternity; human; pedigree;
linkage analysis; genetic disease; animal; plant; breeding; locus;
hybridisation; chromosome; ds.
OS Homo sapiens.
PN US5582979-A.
PD 10-DEC-1996.
PF 21-APR-1989; 341562.
PR 21-APR-1989; US-341562.
PR 05-SEP-1991; US-754351.
PR 04-APR-1994; US-222177.
PA (MARS-) MARSHFIELD CLINIC.
PI Weber JL.
DR WPI; 97-042299/04.
PT Detection of polymorphic genetic markers of the form
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2000, 21:22:20 ; Search time 13789.4 seconds
(without alignments)
8524.491 Million cell updates/sec

Title: US-08-852-495C-2_COPY_168000_197000
Perfect score: 29001
Sequence: 1 ATTACAGATGGAGAAACCA.....GGATTAGGATCATGATCTC 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues
Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	89	0.3	105	105	AQ282107	AQ282107 RPCI11-94
2	88.8	0.3	108	84	B65160	B65160 CIT-HSP-201
3	87.2	0.3	109	94	AQ028426	AQ028426 CIT-HSP-2
4	86.8	0.3	106	97	AQ03692	AQ03692 ag81a10.r
5	85	0.3	101	35	AA583697	AA583697 nn58f10.s
6	85.2	0.3	106	63	AI991750	AI991750 wt48601.x
7	84.8	0.3	107	33	AA385808	AA385808 EST99495
8	85	0.3	109	22	H11143	H11143 ym09c06.r1
9	85.2	0.3	109	24	N25299	N25299 yw52c09.s1
10	85	0.3	110	30	AA244245	AA244245 nc07a04.s
11	85	0.3	110	30	AA244245	AA244245 nc07a04.s
12	84.6	0.3	107	35	AA565533	AA565533 nk42b11.s
13	83.8	0.3	103	108	AQ535244	AQ535244 RPCI-11-3
14	84	0.3	109	103	AQ200347	AQ200347 RPCI11-43
15	83.2	0.3	106	94	AQ046231	AQ046231 RPCI11-36
16	82.8	0.3	102	36	AA654562	AA654562 nt75f10.s
17	82.8	0.3	106	30	AA250812	AA250812 zs06a05.s
18	82.8	0.3	107	24	H67040	H67040 yu68c01.r1
19	82.4	0.3	109	30	AA243009	AA243009 zr25h02.s
20	81.8	0.3	103	108	AQ535244	AQ535244 RPCI-11-3
21	82	0.3	106	105	AQ282340	AQ282340 RPCI11-80
22	81.6	0.3	104	29	AA129957	AA129957 zn68h04.r
23	81.6	0.3	105	30	AA218889	AA218889 zq15d04.s
24	81.8	0.3	109	84	B17434	B17434 345K2.TVB C
25	81.8	0.3	109	94	AQ028426	AQ028426 CIT-HSP-2
26	81.4	0.3	108	84	B65160	B65160 CIT-HSP-201
27	81	0.3	105	28	AA078003	AA078003 7h12d08 C
28	81	0.3	106	44	AI249096	AI249096 qn73g09.x
29	80.6	0.3	103	108	AQ584425	AQ584425 RPCI-11-4
30	80.6	0.3	104	105	AQ321855	AQ321855 RPCI11-11
31	80.6	0.3	106	94	AQ062963	AQ062963 CIT-HSP-2
32	80.8	0.3	108	84	B15423	B15423 345B10.TV C
33	80.2	0.3	101	39	AA835205	AA835205 ak64n01.s
34	80.4	0.3	106	38	AA812141	AA812141 oB48h02.s
35	80.4	0.3	106	106	AQ14071	AQ14071 RPCI-11-1
36	79.8	0.3	101	33	AA381369	AA381369 EST94442
37	80	0.3	104	108	AQ544583	AQ544583 CITBI-El-
38	79.6	0.3	102	84	B48088	B48088 RPCI11-4N6
39	79.8	0.3	107	62	AI933497	AI933497 wn74d02.x
40	79.8	0.3	110	39	AA897366	AA897366 am06h02.s
41	79.4	0.3	105	105	AQ276193	AQ276193 CITBI-El-
42	79.4	0.3	107	24	H67040	H67040 yu68c01.r1
43	79.6	0.3	110	106	AQ386882	AQ386882 RPCI11-13
44	79.6	0.3	110	109	AQ634950	AQ634950 RPCI-11-4
45	79.2	0.3	100	30	AA252633	AA252633 zq43g05.r

ALIGNMENTS

RESULT 1
AQ282107 AQ282107 105 bp DNA GSS 27-APR-1999
LOCUS RPCI11-94B21.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-94B21,
DEFINITION genomic survey sequence.

ACCESSION AQ282107
VERSION AQ282107.1 GI:3907976

KEYWORDS GSS.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 105)

AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE Use of human BAC End Sequences for Sequence-Ready Map Building

JOURNAL Unpublished (1998)

COMMENT Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadamstigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

FEATURES

source

1..105
/organism="Homo sapiens"
/db_xref="GDB:7535756"
/db_xref="taxon:9606"
/clone="RPCI-11-94B21"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT 26 a 31 c 30 g 18 t

Query Match 0.3%; Score 89; DB 105; Length 105;

Best Local Similarity 90.5%; Pred. No. 0.34; Mismatches 0; Gaps 0;

Matches 95; Conservative 0; Indels 10; Indels 0; Gaps 0;

QY 18960 GCTCACCGCTGTATCCAGCACATTTGGGAGGCCAGCGGACAGATCAGAGTTCAGGA 19019

Db 1 GCTCACCGCTGTATCCAGCACATTTGGGAGGCCAGCGGACAGATCAGAGTTCAGGA 60

QY 19020 GTTTGACACCGCTGACCAACATGGTGAACCTGTCTCTACTA 19064

Db 61 GTACGAGACCGCTGACCAACATGGTGAACCTGTCTCTACTA 105

RESULT 2

B65160/c

LOCUS B65160 108 bp DNA GSS 21-JUN-1998

DEFINITION CIT-HSP-2017G2.TRB CIT-HSP Homo sapiens genomic clone 2017G2,

genomic survey sequence.

ACCESSION B65160

VERSION B65160.1 GI:2639138

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 108)

AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,

Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,

Simon,M. and Venter,J.C.

TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map

Building

JOURNAL Unpublished (1997)

COMMENT Other_GSSs: CIT-HSP-2017G2.TPB

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadamstigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html

Seq primer: M13 Reverse

Class: BAC ends.

FEATURES

source

1..108

```

/organism="Homo sapiens"
/db_xref="GDB:7043860"
/db_xref="taxon:9606"
/clone="201762"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/notes="Vector: pBeloBAC11; Site_1: HindIII; Site_2: HindIII"
BASE COUNT      26 a      27 c      34 g      21 t
ORIGIN

Query Match      0.3%; Score 88.8; DB 84; Length 108;
Best Local Similarity 88.9%; Pred. No. 0.36;
Matches 96; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 24883 AGAATGCGGTGTGTGCTTGGCTCACTGCAACCTCCGCTCTGGGTCAAGCAATTC 24942
|| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 108 AGTGTGCAGTGTGTATGATCTTGGCTCACTGCAACCTCCGCTCTGGGTCAAGATTC 49

QY 24943 CCATGCTCAGCTCCCGAGTGGGATTACAGGCACATGCCACCA 24990
| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 48 TCCTGCTCAGCTCTCTGAGTAGCTGGGATTACAGGCGCATGCCACCA 1

RESULT 3
LOCUS      AQ028426      109 bp      DNA      GSS      30-JUN-1998
DEFINITION CIT-HSP-2313G15.TF CIT-HSP Homo sapiens genomic clone 2313G15,
genomic survey sequence.
ACCESSION      AQ028426
VERSION      AQ028426.1 GI:3268648
KEYWORDS      GSS.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 109)
AUTHORS      Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE      Use of a random BAC End Sequence Database for Sequence-Ready Map
BUILDING      Building (1998)
JOURNAL      Unpublished (1998)
COMMENT      Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..109
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2313G15"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT      19 a      36 c      25 g      29 t
ORIGIN

Query Match      0.3%; Score 87.2; DB 94; Length 109;
Best Local Similarity 88.0%; Pred. No. 0.54;
Matches 94; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 18963 CACGGCTTAATCCAGCAGCTTTGGAGCCGAGCGGCACATCAGAGGTCAGAGTT 19022
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 106 CACGCGTGTAAATCCAGCAGCTTTGGAGGCTGAGCGGCACATCAGAGGTCAGAGAT 47

QY 19023 TCAGACCAAGCCTTGACCAACATGTTGAAACCTCTCTCTACTAACAA 19068
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 46 CGAGACCATCTCTGGCTTAACACGGTGAACCTCTCTCTACTAAAAA 1

Matches 95; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 24586 TTTTTCCTGAAATGAGTCTCACTCTGTCGCCAGCTGGAGTACAGTGGCACAATCTTG 24645
| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 TGTTCCTGAGAGGAGCTCTCACTCTGTCACCCAGCTGGAGTGGCAGGTCAGTCTGA 61

QY 24646 GTTCTACTCAACCTCCACCTCCTGGGTTCAAGCGAGTCTCTCTGACTCA 24693
| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 62 GCTCACTCAACCTCCACCTCCTGGGTTCAAGCGATTCTCTCTGCTCA 109

RESULT 4
LOCUS      AA703692/c      106 bp      mRNA      EST      24-DEC-1997
DEFINITION ag1a10.r1 Stratagene hMT neuron (#937233) Homo sapiens cDNA clone
IMAGE:1140858 5' similar to contains Alu repetitive element,, mRNA
sequence.
ACCESSION      AA703692
VERSION      AA703692.1 GI:2713610
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 106)
AUTHORS      Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wyllie,T., Waterston,K. and Wilson,R.
TITLE      WashU-NCI human EST project
JOURNAL      Unpublished (1997)
COMMENT      On Sep 12, 1996 this sequence version replaced gi:1397630.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LUNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28ml3 revl Et from Amersham
High quality sequence stop: 53.
FEATURES             Location/Qualifiers
     source           1..106
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1140858"
                     /clone_lib="Stratagene hMT neuron (#937233)"
                     /dev_stage="hMT neurons"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally. Primer: Oligo dT.
                     Differentiated, post mitotic hMT neurons. Average insert
                     size: 1.5 Kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'
                     GAATTCGCGCAGAG 3' -3' adaptor sequence: 5'
                     CTCGAGTTTCTTTTCTTTT 3'"
BASE COUNT      19 a      29 c      29 g      29 t
ORIGIN

Query Match      0.3%; Score 86.8; DB 37; Length 106;
Best Local Similarity 88.7%; Pred. No. 0.61;
Matches 94; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 18963 CACGGCTTAATCCAGCAGCTTTGGAGCCGAGCGGCACATCAGAGGTCAGAGTT 19022
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 106 CACGCGTGTAAATCCAGCAGCTTTGGAGGCTGAGCGGCACATCAGAGGTCAGAGAT 47

QY 19023 TCAGACCAAGCCTTGACCAACATGTTGAAACCTCTCTCTACTAACAA 19068
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 46 CGAGACCATCTCTGGCTTAACACGGTGAACCTCTCTCTACTAAAAA 1

```

```

RESULT 5
AA583697/c
LOCUS
DEFINITION nm58f10.s1 NCI_CGAP_Kid6 Homo sapiens cDNA clone IMAGE:1088107 3'
similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION AA583697
VERSION AA583697.1 GI:2368306
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 101)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Apr 14, 1993 this sequence version replaced gi:692704.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Stratagene, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1890 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 93.
FEATURES
Location/Qualifiers
source
1..101
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1088107"
/clone_lib="NCI_CGAP_Kid6"
/sex="mixed"
/tissue_type="kidney tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: kidney; Vector: Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Pooled kidney tumors. 5' adaptor sequence: 5'
GAATTCGGCAGAG 3' 3' adaptor sequence: 5'
CTCAGATTTTTTTTTTTTTT 3' Average insert size: 1.0 kb."
BASE COUNT 21 a 29 c 28 g 23 t
ORIGIN

Query Match 0.3%; Score 85; DB 35; Length 101;
Best Local Similarity 90.18; Pred. No. 0.98;
Matches 91; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 18962 TCACGCCGTGTAATCCAGCACTTTGGAGGCCAGCGGAGATCAGCAGGTCTGAGT 19021
Db 101 TCACGCCGTGTAATCCAGCACTTTGGAGGCCAGCGGAGATCAGCAGGTCTGAGT 42

Qy 19022 TTGAGCAGCGCTGACCAACATGGTGAACCTGTCTCTAC 19062
Db 41 TCGAGATCAGCGCTGGCCCAACATGGTGAACCTGTCTCTAC 1

RESULT 6
AI991750/c
LOCUS
DEFINITION wt48e01.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2510712 3'
similar to contains Alu repetitive element;contains element LTR8
repetitive element ;, mRNA sequence.
ACCESSION AI991750
VERSION AI991750.1 GI:5838578
KEYWORDS EST.

```

```

human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 106)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Dec 20, 1995 this sequence version replaced gi:1133359.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 62.
FEATURES
Location/Qualifiers
source
1..106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2510712"
/clone_lib="NCI_CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"
BASE COUNT 24 a 23 c 22 g 37 t
ORIGIN

Query Match 0.3%; Score 85.2; DB 63; Length 106;
Best Local Similarity 87.7%; Pred. No. 0.91;
Matches 93; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 3343 AATCTAGCACTTTGGGAGCCGAGCGGAGATCAGCTAGGTCTGAGGACC 3402
Db 106 AATCCAGCACTTTGGGAGCGTGGTGTAGATCAGCTAGGTCTGAGGACT 47

Qy 3403 AGCTGACCAACATGGCAAAACCTCATCTACTACAAAATACAAA 3448
Db 46 AACCTGCCAACATGGTAAACCTCATCTATTAAAAA 1

RESULT 7
AA385808/c
LOCUS
DEFINITION AA385808 107 bp mRNA EST 21-APR-1997
EST99495 Thyroid Homo sapiens cDNA 5' end similar to EST containing
Alu repeat, mRNA sequence.
ACCESSION AA385808
VERSION AA385808.1 GI:2038127
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 107)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Lily,

```

Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wang, J., Xu, C., Yu, G.D., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

TITLE
Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

JOURNAL
Nature 377 (6547 Suppl), 3-174 (1995)

MEDLINE
12140200

COMMENT
On Jan 25, 1995 this sequence version replaced gi:637865.
Contact: Kerlavage, AR

Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423

Email: arkerlavet@tigr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tadb/hgi/hgi.html>)
Seq primer: M13 Reverse.

FEATURES

source
Location/Qualifiers
1..107
/organism="Homo sapiens"
/db_xref="ATCC (inhost):189984"
/db_xref="taxon:9606"
/clone_lib="Thyroid"
/dev_stage="adult"
/note="Organ: thyroid gland; Vector: pBluescript SK-;
Site_1: ECORI; Site_2: XhoI"
16 a 34 c 28 g 26 t 3 others

Query Match 0.3%; Score 84.8; DB 33; Length 107;
Best Local Similarity 86.0%; Pred. No. 1;
Matches 92; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 19121 GGGACACTGAGCAGAGAGATCACTTGACCCAGGAGGAGAGATTGCGAGCTGAGA 19180

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

DB 107 GGGAGCTGAGCAGGAGAGAACGCTGTAACCCAGGAGGAGCTTGCAGTGCAGCTGAGA 48

QY 19181 TCGCGCACTGCATCCACGCTGGGAGACAGCGAGACTCCATCTC 19227

DB 47 TCTCGCCACTGCATCCACCTGGGAGGAGGAGGAGACTGCTCTC 1

RESULT

8
LOCUS H11143 109 bp mRNA EST 26-JUN-1995
DEFINITION Ym09c06.r1 Soares infant brain lntb Homo sapiens cDNA clone
IMAGE:47310 5' similar to contains Alu repetitive element; contains
MER22 repetitive element ;, mRNA sequence.

ACCESSION H11143

VERSION H11143.1 GI:875963

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 109)

AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,

Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marr, M.,

Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,

Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and

Wilson, R.

The WashU-Merck EST Project

Unpublished (1995)

TITLE
On May 5, 1995 this sequence version replaced gi:798506.
Contact: Wilson RK
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
Insert Size: 1316

Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Putative full length read

Insert Length: 1316 Std Error: 0.00

Seq primer: M13RP1

High quality sequence stop: 363.

FEATURES

source

Location/Qualifiers

1..109

/organism="Homo sapiens"

/db_xref="GDB:419851"

/db_xref="taxon:9606"

/clone="IMAGE:47310"

/clone_lib="Soares infant brain lntb"

/sex="female"

/dev_stage="73 days post natal"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: whole brain; Vector: Lfamid BA; Site_1: Not

I; Site_2: Hind III; 1st strand cDNA was primed with a Not

I - oligo(dT) primer [5,

AACTGGAGAATCGCGCGCAGCAATTTTTTTTTTTT 3'];

double-stranded cDNA was ligated to Hind III adaptors

(Pharmacia), digested with Not I and directionally cloned

into the Not I and Hind III sites of the Lfamid BA vector.

Library went through one round of normalization. Library

constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 27 a 28 c 36 g 18 t

ORIGIN

Query Match 0.3%; Score 85; DB 22; Length 109;

Best Local Similarity 86.2%; Pred. No. 0.95;

Matches 94; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 2499 TGAGACGGAGTCTCACTCTGTCACCAGGCTGGAGTGAGTACTCGGCTCACT 2558

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

DB 109 TGAGAAGGCGTCTCACTCTGTCACCAGGCTGGAGTGAGTCTCACTCACTCACT 50

QY 2559 GCACCTCCGCCCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCCCGA 2607

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

DB 49 GGAACCTCGCTCCCGGTTCAAGTGATTCTCTGCTTAGCTCTCTGA 1

RESULT 9

N25299/c

LOCUS N25299 109 bp mRNA EST 28-DEC-1995

DEFINITION YW52c09.s1 Weizmann Olfactory Epithelium Homo sapiens cDNA clone

IMAGE:255856 3' similar to contains Alu repetitive element;; mRNA

sequence.

ACCESSION N25299

VERSION N25299.1 GI:1139449

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 109)

AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,

Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,

Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,

Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,

Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierly-Meg, J.,

Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.

and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags

Genome Res. 6 (9), 807-828 (1996)

TITLE
JOURNAL 9704478

MEDLINE
9704478

COMMENT
On Apr 14, 1993 this sequence version replaced gi:837394.

High quality sequence stop: 90.

FEATURES

source
1. .110
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1007406"
/clone_lib="NCI_CGAP_Prl"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected, histologically normal prostate epithelial cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."
BASE COUNT 17 a 26 c 28 g 38 t 1 others
ORIGIN

Query Match 0.3%; Score 85; DB 30; Length 110;
Best Local Similarity 85.5%; Pred. No. 0.95;
Matches 94; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 19128 TGAGGCGAGAGATCACTTGAACCCAGGAGGAGAGATTGCGAGCTGAGATCGCGCC 19187
Db 110 TGAGGCGAGAGATCTTTGAACCCAGGAGGAGAGATTGCGAGCTGAGCAAGATCTGCC 51
QY 19188 ACTGCATTCAGCCCTGGGAGACAGCGAGAGCTCCATCTCAAAATTA 19237
Db 50 ACTGCATCCAGCCTGGGCAACAGATCAAGACTCCATCTCAAAAAA 1

RESULT 12
AA565533/c 107 bp mRNA EST 08-SEP-1997
LOCUS nk42b11.s1 NCI_CGAP_GC2 Homo sapiens cDNA clone IMAGE:1016157 3'
DEFINITION similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION AA565533
VERSION AA565533.1 GI:2337172
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 107)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1393355.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1350
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc., David B. Krizman, Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 1661 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 87.
Location/Qualifiers

source

1. .107
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1016157"
/clone_lib="NCI_CGAP_GC2"
/tissue_type="germ cell tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Bulk germ cell tumor. 5' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.2 kb."
BASE COUNT 22 a 34 c 26 g 25 t
ORIGIN

Query Match 0.3%; Score 84.6; DB 35; Length 107;
Best Local Similarity 86.9%; Pred. No. 1.1;
Matches 93; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 19093 TGGCATCTCCCTAGTCCAGCTACTCGGACACTGAGCGAGAGAACTTGAACCC 19152
Db 107 TGTGTGTGCTGTGAATCCAGCTACTCAGGAGGCTGAGCGAGAACTTGAACCT 48

QY 19153 AGGAGCGAGAGATTGCGAGCTGAGATCGCGCCACTGCATTCCAG 19199
Db 47 GGGAGCGAGAGCTTGCAGTGCAGCTGAGATTGAGCCACTGCATCTCCAG 1

RESULT 13
AA535244 103 bp DNA GSS 18-MAY-1999
LOCUS RPCI-11-317H22.TV RPCI-11 Homo sapiens genomic clone
DEFINITION RPCI-11-317H22, genomic survey sequence.
ACCESSION AA535244
VERSION AA535244.1 GI:4846934
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@edejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html. Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1. .103
/organism="Homo sapiens"
/db_xref="GDB:762153"
/db_xref="taxon:9606"
/clone="RPCI-11-317H22"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"

FEATURES

source

```
BASE COUNT      31 a      27 c      27 g      18 t
ORIGIN

Query Match      0.3%; Score 83.8; DB 108; Length 103;
Best Local Similarity 88.3%; Pred. No. 1.3;
Matches 91; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 19961 CCAGCATTGTTGGAGCGGCGTGCAGATCCTTGTAGGACAGGAGTTCAAGACGAGCC 20020
|||||
1 CCAGCATTGTTGGAGCGGCGGAGCGGCGAGATCCTTGTAGGTCAGGAGTTCCAGACGAGCC 60
|||||

Qy 20021 TGGCTAACATGCAAAACCCCATCTCTACTATAAAATACAAAA 20063
|||||
61 TGGCCAACATGTTGTAACCCCGTCTCTGTCTATATAATACAAAA 103
|||||

RESULT 14
LOCUS      AQ200347      109 bp      DNA      GSS      20-APR-1999
DEFINITION      RPC111-43B21..TJ RPCI-11 Homo sapiens genomic clone RPCI-11-43B21,
ACCESSION      AQ200347
VERSION      AQ200347.1 GI:3612546
KEYWORDS      GSS.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      1 (bases 1 to 109)
AUTHORS      Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
            Venter,J.C.
TITLE      Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL      Unpublished (1998)
COMMENT      Other_GSSs: RPC111-43B21.TK
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
            Class: BAC ends.

FEATURES             Location/Qualifiers
     source           1..109
                     /organism="Homo sapiens"
                     /db_xref="GDB:7516172"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-43B21"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPC111 Human Male BAC Library"

BASE COUNT      27 a      31 c      29 g      22 t
ORIGIN

Query Match      0.3%; Score 84; DB 103; Length 109;
Best Local Similarity 86.1%; Pred. No. 1.2;
Matches 93; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 20012 AGACCAGCTGGCTAACATGGCAAAACCCCATCTCTACTATAAAATACAAAAATTAACGAC 20071
|||||
2 ATACCAGCTGGCGACACGCGTGAACCCCATCTCTACTTAAGAGTACATTAATAGCGAG 61
|||||

Qy 20072 GCGTGGTGGTGACGCCCTGTATATCCAGCTACTCTGGAGGCTGAGGCA 20119
|||||

BASE COUNT      31 a      27 c      27 g      18 t
ORIGIN

Query Match      0.3%; Score 83.2; DB 94; Length 106;
Best Local Similarity 87.5%; Pred. No. 1.5;
Matches 91; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 27795 CTCTGCCCTCCTGTTTAAAGCGATTCTCTGCCCTCAGACTCCTGAGTAGCTGAGATTACA 27854
|||||
105 CTCCTCCTACCGTGTTTTAGCGATTCTCATGCCCTCAGCTCTTGTAGTAGTGGGACTATA 46
|||||

Qy 27855 GGTGCCTGCCACCATGCCAGCTAATTTTATATTTTAGTAGA 27898
|||||
45 GGTGTCTGCCACCATGCCAGCTAATTTTGTATTTTAGTAGA 2
|||||

Search completed: June 20, 2000, 09:51:12
Job time: 508439 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2000, 02:28:19 ; Search time 599.42 seconds
(without alignments)
6288.907 Million cell updates/sec

Title: US-08-852-495C-2_COPY_168000_197000
Perfect score: 23001
Sequence: 1 ATTTACAGATGGAGAACCA.....GGATTAGGATCATGATCTC 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	82.6	0.3	105	4	US-08-481-658B-65 Sequence 65, Appl
C 2	82.6	0.3	105	4	US-08-477-504A-65 Sequence 65, Appl
C 3	82.6	0.3	105	4	US-08-486-756A-65 Sequence 65, Appl
C 4	82.6	0.3	105	4	US-08-485-862B-65 Sequence 65, Appl
C 5	82.6	0.3	105	5	US-08-787-739-65 Sequence 65, Appl
C 6	68.6	0.2	84	3	US-08-454-557C-91 Sequence 91, Appl
C 7	68.6	0.2	84	4	US-08-340-426D-91 Sequence 91, Appl
C 8	68.6	0.2	84	4	US-08-450-673C-91 Sequence 91, Appl
C 9	68.6	0.2	84	6	PCT-US95-17111A-91 Sequence 91, Appl
C 10	65.6	0.2	105	4	US-08-481-658B-65 Sequence 65, Appl
C 11	65.6	0.2	105	4	US-08-477-504A-65 Sequence 65, Appl
C 12	65.6	0.2	105	4	US-08-486-756A-65 Sequence 65, Appl
C 13	65.6	0.2	105	4	US-08-485-862B-65 Sequence 65, Appl
C 14	65.6	0.2	105	5	US-08-787-739-65 Sequence 65, Appl
C 15	64	0.2	92	1	US-08-222-177A-430 Sequence 430, App
C 16	59.4	0.2	78	3	US-08-454-557C-70 Sequence 70, Appl
C 17	59.4	0.2	78	4	US-08-340-426D-70 Sequence 70, Appl
C 18	59.4	0.2	78	4	US-08-450-673C-70 Sequence 70, Appl
C 19	59.4	0.2	78	6	PCT-US95-17111A-70 Sequence 70, Appl
C 20	59	0.2	60	1	US-08-222-177A-244 Sequence 244, App
C 21	57.8	0.2	78	3	US-08-454-557C-70 Sequence 70, Appl
C 22	57.8	0.2	78	4	US-08-340-426D-70 Sequence 70, Appl
C 23	57.8	0.2	78	4	US-08-450-673C-70 Sequence 70, Appl
C 24	57.8	0.2	78	6	PCT-US95-17111A-70 Sequence 70, Appl
C 25	56	0.2	85	3	US-08-454-557C-92 Sequence 92, Appl
C 26	56	0.2	85	4	US-08-340-426D-92 Sequence 92, Appl
C 27	56	0.2	85	4	US-08-450-673C-92 Sequence 92, Appl

C	28	56	0.2	85	6	PCT-US95-17111A-92	Sequence 92, Appl
	29	55	0.2	84	3	US-08-454-557C-91	Sequence 91, Appl
	30	55	0.2	84	4	US-08-340-426D-91	Sequence 91, Appl
	31	55	0.2	84	4	US-08-450-673C-91	Sequence 91, Appl
	32	55	0.2	84	6	PCT-US95-17111A-91	Sequence 91, Appl
	33	53.6	0.2	91	1	US-08-222-177A-166	Sequence 166, App
	34	51.2	0.2	56	1	US-08-222-177A-65	Sequence 65, Appl
	35	50.4	0.2	60	3	US-08-454-557C-60	Sequence 60, Appl
	36	50.4	0.2	60	4	US-08-340-426D-60	Sequence 60, Appl
	37	50.4	0.2	60	4	US-08-450-673C-60	Sequence 60, Appl
	38	50.4	0.2	60	6	PCT-US95-17111A-60	Sequence 60, Appl
	39	50	0.2	76	3	US-08-454-557C-69	Sequence 69, Appl
	40	50	0.2	76	4	US-08-340-426D-69	Sequence 69, Appl
	41	50	0.2	76	4	US-08-450-673C-69	Sequence 69, Appl
	42	50	0.2	76	6	PCT-US95-17111A-69	Sequence 69, Appl
	43	50	0.2	83	4	US-08-481-658B-66	Sequence 66, Appl
	44	50	0.2	83	4	US-08-477-504A-66	Sequence 66, Appl
	45	50	0.2	83	4	US-08-486-756A-66	Sequence 66, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-65/c
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/COCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.3%; Score 82.6; DB 4; Length 105;
Best Local Similarity 86.7%; Pred. No. 2.5e-08;

Matches 91; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 18973 ATCCGAGCATTGGAGCGCAAGCGACAGATCAGAGTCTAGGAGTTTGAGACGAGC 19032
|||||
Db 105 ATCCGAGCATTGGAGCGCGAGCTGGTGATCACAAGTCTAGGAGTTTGAGACGAGC 46
|||||
QY 19033 CTGACCAACATGGTGAACCCCTGTCTTACTACTAACAATAACAAA 19077
|||||
Db 45 CTGGCCATATGGTGAACCCCTGTCTTACTTAAAGATGTAATAAAA 1

RESULT 2

US-08-477-504A-65/c
; Sequence 65, Application US/08477504A
; Patent No. 5972353
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,504A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-477-504A-65

Query Match 0.3%; Score 82.6; DB 4; Length 105;
Best Local Similarity 86.7%; Pred. No. 2.5e-08;
Matches 91; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 18973 ATCCGAGCATTGGAGCGCAAGCGACAGATCAGAGTCTAGGAGTTTGAGACGAGC 19032
|||||
Db 105 ATCCGAGCATTGGAGCGCGAGCTGGTGATCACAAGTCTAGGAGTTTGAGACGAGC 46
|||||
QY 19033 CTGACCAACATGGTGAACCCCTGTCTTACTACTAACAATAACAAA 19077
|||||
Db 45 CTGGCCATATGGTGAACCCCTGTCTTACTTAAAGATGTAATAAAA 1

RESULT 3

US-08-486-756A-65/c
; Sequence 65, Application US/08486756A
; Patent No. 5981711
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,756A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-486-756A-65

Query Match 0.3%; Score 82.6; DB 4; Length 105;
Best Local Similarity 86.7%; Pred. No. 2.5e-08;
Matches 91; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 18973 ATCCGAGCATTGGAGCGCAAGCGACAGATCAGAGTCTAGGAGTTTGAGACGAGC 19032
|||||
Db 105 ATCCGAGCATTGGAGCGCGAGCTGGTGATCACAAGTCTAGGAGTTTGAGACGAGC 46
|||||
QY 19033 CTGACCAACATGGTGAACCCCTGTCTTACTACTAACAATAACAAA 19077
|||||
Db 45 CTGGCCATATGGTGAACCCCTGTCTTACTTAAAGATGTAATAAAA 1

RESULT 4

US-08-485-862B-65/c
; Sequence 65, Application US/08485862B
; Patent No. 5989838
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court

```
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485.862B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-485-862B-65

Query Match 0.3%; Score 82.6; DB 4; Length 105;
Best Local Similarity 86.7%; Pred No. 2.5e-08;
Matches 91; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 18973 ATCCGAGCACTTTGGGAGCGCCGAGATCACGAGGTTCAGGAGTTTGAGACCAGC 19032
Db 105 ATCCGAGCACTTTGGGAGCGCGAGCTGTGGATCACAGGTTCAGGAGTTTGAGACCAGC 46

QY 19033 CTGACCAACATGTTGAAACCCCTGTCTCTACTACAAAATACAAAA 19077
Db 45 CTGGCCAATATGTTGAAACCCCTGTCTCTACTAAAGATGTAAAAA 1

RESULT 5
US-08-787-739-65/C
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-787-739-65

Query Match 0.3%; Score 82.6; DB 5; Length 105;
Best Local Similarity 86.7%; Pred. No. 2.5e-08;
Matches 91; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 18973 ATCCGAGCACTTTGGGAGCGCCGAGATCACGAGGTTCAGGAGTTTGAGACCAGC 19032
Db 105 ATCCGAGCACTTTGGGAGCGCGAGCTGTGGATCACAGGTTCAGGAGTTTGAGACCAGC 46

QY 19033 CTGACCAACATGTTGAAACCCCTGTCTCTACTACAAAATACAAAA 19077
Db 45 CTGGCCAATATGTTGAAACCCCTGTCTCTACTAAAGATGTAAAAA 1

RESULT 6
US-08-454-557C-91/C
; Sequence 91, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; US-08-454-557C-91

Query Match 0.2%; Score 68.6; DB 3; Length 84;
Best Local Similarity 89.2%; Pred. No. 1.5e-05;
Matches 74; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 18963 CACGCTGTATCCAGCACCTTTGGGAGGCCAAGCGGACAGATCAGAGTCTAGGAGTT 19022
Db 83 CACGCTGTATCCAGCACCTTTGGGAGGCCAAGCGGAGGCTGAGCGGGCGGATCAGAGTCTAGGAGTT 24

Qy 19023 TGAGACCAGCCTGACCAACATGG 19045
Db 23 CGACACCAGCCTGATGACATGG 1

RESULT 7
US-08-426D-91/c
; Sequence 91, Application US/08340426D
; Patent No. 5948634
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; US-08-340-426D-91

Query Match 0.2%; Score 68.6; DB 4; Length 84;
Best Local Similarity 89.2%; Pred. No. 1.5e-05;
Matches 74; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 18963 CACGCTGTATCCAGCACCTTTGGGAGGCCAAGCGGACAGATCAGAGTCTAGGAGTT 19022
Db 83 CACGCTGTATCCAGCACCTTTGGGAGGCCAAGCGGAGGCTGAGCGGGCGGATCAGAGTCTAGGAGTT 24

Qy 19023 TGAGACCAGCCTGACCAACATGG 19045
Db 23 CGACACCAGCCTGATGACATGG 1

RESULT 8
US-08-450-673C-91/c
; Sequence 91, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; US-08-450-673C-91

Query Match 0.2%; Score 68.6; DB 4; Length 84;
Best Local Similarity 89.2%; Pred. No. 1.5e-05;
Matches 74; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 18963 CACGCTGTATCCAGCACCTTTGGGAGGCCAAGCGGACAGATCAGAGTCTAGGAGTT 19022
Db 83 CACGCTGTATCCAGCACCTTTGGGAGGCCAAGCGGAGGCTGAGCGGGCGGATCAGAGTCTAGGAGTT 24

Qy 19023 TGAGACCAGCCTGACCAACATGG 19045
Db 23 CGACACCAGCCTGATGACATGG 1

```
RESULT 9
PCT-US95-17111A-91/c
; Sequence 91, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/17111A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,426
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
PCT-US95-17111A-91

Query Match 0.2%; Score 68.6; DB 6; Length 84;
Best Local Similarity 89.2%; Pred. No. 1.5e-05;
Matches 74; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 18963 CACGCTGTAATCCAGCACCTTTGGAGGCGCAAGCGGACAGATCACGAGGTGAGGTT 19022
Db 83 CACGCTGTAATCCAGCACCTTTGGAGGCTGAGGCGGCGGATCACGAGGTGAGGTT 24

Qy 19023 TGAGACACCGCTGACCAACATGG 19045
Db 23 CGACACCGCTGATGAACATGG 1

RESULT 10
US-08-481-658B-65
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,504A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
```

```
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.2%; Score 65.6; DB 4; Length 105;
Best Local Similarity 84.3%; Pred. No. 6.9e-05;
Matches 86; Conservative 0; Mismatches 14; Indels 2; Gaps 1;

Qy 27880 TTTTATATTTTAGTAGAGACGGGTTTCAGCATGTTGGTCAAGCTTGTCTCAAACTCC 27939
Db 2 TTTTACATCTTAGTAGAGACGGTTTCACCATATTTGGCCAGGCTGCTCTCAAACTCC 61

Qy 27940 TGACCTCAGGTATCCACCCACCTCGGCTCCGAAAGTGTG 27981
Db 62 TGACCT--TGTATCCACGAGCTCGGCTCCCAAGTGTG 101

RESULT 11
US-08-477-504A-65
; Sequence 65, Application US/08477504A
; Patent No. 5972353
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,504A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
```

Query Match 0.28; Score 65.6; DB 4; Length 105;

Db 80 AACACACACACA 91

Search completed: June 20, 2000, 18:07:04
Job time: 537154 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2000, 17:56:59 ; Search time 29135.9 Seconds
(without alignments)
-601.018 Million cell updates/sec

Title: US-08-852-495C-2_COPY_196000_214000
 Perfect score: 18001
 Sequence: 1 CTAAGAAGTATTTTAAAC...TTGGACATGCTGTGTTCCCTTC 18001

Scoring table: IDENTITY__NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

```
Minimum DB seq length: 10
Maximum DB seq length: 110
```

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl : ★

1: gb_ba1:*
2: gb_ba2:*
3: gb_em:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_on:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_hg1:*
33: gb_hg2:*
34: gb_in1:*
35: gb_in2:*
36: gb_in3:*
37: gb_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_hg3:*
42: gb_hg4:*
43: gb_hg5:*
44: gb_hg6:*

45: gb_hgt7:*
46: em_hgt1:*
47: em_hgt2:*
48: em_hgt3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_hgt8:*
53: gb_hgt9:*
54: gb_hgt10:*
55: gb_hgt11:*
56: gb_hgt12:*
57: gb_hgt13:*
58: gb_hgt14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	88.8	0.5	108	10	HSLDLRN2	X05250 Human LDL-r
C 2	87.4	0.5	108	11	HSU67803	U67803 Human small
C 3	84.4	0.5	103	9	HUMALCE221	M87896 Human carci
C 4	84	0.5	108	10	HSLDLRN2	X05250 Human LDL-r
C 5	83	0.5	108	10	HSLDLI12	X05248 Human LDL-r
C 6	82.2	0.5	107	9	HUMALCE162	M87924 Human carci
C 7	79.8	0.4	108	10	HSLDLRD1	X05249 Human LDL-r
C 8	79.8	0.4	108	10	HSLDLRD1	X05249 Human LDL-r
C 9	79.8	0.4	108	10	HSLDLRD2	X05251 Human LDL-r
C 10	79.8	0.4	108	10	HSLDLRD2	X05251 Human LDL-r
C 11	79.2	0.4	103	13	HS81C8R	X57789 Human sequ
C 12	77.8	0.4	108	11	HSU67804	U67804 Human small
C 13	77.2	0.4	91	13	HUMUT8164A	L30244 Human STS U
C 14	77.2	0.4	110	11	HSU67807	U67807 Human small
C 15	76.2	0.4	108	11	HSU67808	U67808 Human small
C 16	75.6	0.4	103	13	HS81C8R	X57789 Human sequ
C 17	74.4	0.4	110	9	HUMALCE43	M87900 Human carci
C 18	74.2	0.4	107	9	HUMALCE162	M87924 Human carci
C 19	74.2	0.4	108	9	HUMD1D03M5	D16965 Human HepG2
C 20	74	0.4	103	9	HUMALCE221	M87896 Human carci
C 21	73.4	0.4	110	11	HSU67807	U67807 Human small
C 22	73.2	0.4	104	9	HUMALCE272	M87899 Human carci
C 23	72.4	0.4	108	10	HSLDLI12	X05248 Human LDL-r
C 24	72	0.4	90	9	HUMD1DLRFL	K03555 Human low d
C 25	72	0.4	106	13	G32743	K03555 Human low d
C 26	71.2	0.4	108	13	G43535	G43535 WIATF-2393-S
C 27	70.4	0.4	97	9	HUMD1DLR2	M14180 Human low d
C 28	70.4	0.4	106	13	G32743	G32743 A009P31 Hum
C 29	69.8	0.4	90	9	HUMD1DLRFL	K03555 Human low d
C 30	69.8	0.4	104	9	HUMALCE272	M87899 Human carci
C 31	69.4	0.4	101	10	S79560	S79560 HRX (intron
C 32	69.2	0.4	97	9	HUMD1DLR1	M14178 Human low d
C 33	69	0.4	108	11	HSU67803	U67803 Human small
C 34	68.8	0.4	99	13	HUMUT7692A	L30306 Human STS U
C 35	68.2	0.4	107	11	HSU67806	U67806 Human small
C 36	68	0.4	108	13	G43535	G43535 WIATF-2393-S
C 37	67.8	0.4	108	3	AF185109S1	AF185109 Lactofl
C 38	67.8	0.4	108	9	HUMD1D03M5	D16965 Human HepG2
C 39	67.6	0.4	100	13	HUMUT931A	L31299 Human STS U
C 40	67.2	0.4	80	9	HUMBRKFAE	M36135 Human alpha
C 41	67.2	0.4	97	9	HUMD1DLR2	M14180 Human low d
C 42	67	0.4	91	13	HUMUT8164A	L30244 Human STS U
C 43	67	0.4	97	9	HUMD1DLR1	M14178 Human low d
C 44	66.8	0.4	94	9	HUMHGAL	M13479 Human alpha
C 45	66.8	0.4	100	11	HSU67848	U67848 Human beta-

ALIGNMENTS

```
RESULT 1
HSLDLRN2/c
LOCUS      108 bp      DNA      PRI      20-MAY-1992
DEFINITION Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION  X05250
VERSION     X05250.1 GI:34337
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 108)
AUTHORS     Williamson, R., Beisiegel, U., Dunning, A., Havinga, J.R.,
            Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
            Eur. J. Biochem. 164 (1), 77-81 (1987)
JOURNAL     87161901
MEDLINE
COMMENT     See X05252 for deletion junction
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES    Location/Qualifiers
            source          1..108
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
            intron          1..108
                        /note="intron XIV fragment"
BASE COUNT  28 a 23 c 39 g 18 t
ORIGIN

Query Match      0.5%; Score 88.8; DB 10; Length 108;
Best Local Similarity 88.9%; Pred. No. 1.3e-05;
Matches 96; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 15729 CTGGCTCATTCGCAACCTCTCCCTCCGGTTCAAGTGATTCCTGACTCAGCTCCCG 15788
            |||||||GCAACCTCTCCCTCCGGTTCAAGTGATTCCTGACTCAGCTCCCG |||||
Db 108 CTGGCTCATTCGCAACCTCTCCCTCCGGTTCAAGTGATTCCTGACTCAGCTCCCG 49

Qy 15789 AGTAGCTGGGATTACAGCATCATCACCATGCTGGGTAATTTTGT 15836
            |||||||GCAACCTCTCCCTCCGGTTCAAGTGATTCCTGACTCAGCTCCCG |||||
Db 48 AGTAGCTGGGATTACAGCATCATCACCATGCTGGGTAATTTTGT 1

RESULT 2
HSU67803/c
LOCUS      108 bp      RNA      PRI      01-AUG-1997
DEFINITION Human small cytoplasmic Alu transcript.
ACCESSION  U67803
VERSION     U67803.1 GI:2289917
KEYWORDS    Alu.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 108)
AUTHORS     Shaikh, T.H., Roy, A.M., Kim, J., Batzer, M.A. and Deininger, P.L.
            cDNAs derived from primary and small cytoplasmic Alu (sAlu)
            transcripts
            J. Mol. Biol. 271 (2), 222-234 (1997)
JOURNAL     97415756
MEDLINE
REFERENCE   2 (bases 1 to 108)
AUTHORS     Shaikh, T.H., Kim, J., Batzer, M.A. and Deininger, P.L.
            Direct Submission
            Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
            Children's Hospital of Philadelphia, 1004F Abramson Research
            Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
            Location/Qualifiers
            source          1..108
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="TscAlu2"
```

```
repeat_region 1..108
            /note="sAlu"
            /rpt_family="Alu"
            /rpt_type="dispersed"
BASE COUNT  23 a 39 c 30 g 16 t
ORIGIN

Query Match      0.5%; Score 87.4; DB 11; Length 108;
Best Local Similarity 93.8%; Pred. No. 2.3e-05;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 15143 GTAGAGATGGAGTTTCGCCGTGTAGCCAGGAGTGTCGATCTCCGACTCGTGATCC 15202
            |||||||GTTTCACCTTGTGTAGCCAGGAGTGTCGATCTCCGACTCGTGATCC |||||
Db 97 GTAGAGACGGGGTTTCACCTTGTGTAGCCAGGAGTGTCGATCTCCGACTCGTGATCC 38

Qy 15203 ACCGGCTCCGGCTCCCAAGAGTCTGGGATTACAGGC 15239
            |||||||GTTTCACCTTGTGTAGCCAGGAGTGTCGATCTCCGACTCGTGATCC |||||
Db 37 GCCCGCTCCGGCTCCCAAGAGTCTGGGATTACAGGC 1

RESULT 3
HUMALCE221/c
LOCUS      103 bp      ss-RNA      PRI      15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION  M87896
VERSION     M87896.1 GI:174874
KEYWORDS    Alu repeat.
SOURCE      Homo sapiens male embryo carcinoma CDNA to other RNA.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 103)
AUTHORS     Sinnett, D., Richer, C., Deragon, J.-M. and Labuda, D.
            Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
            J. Mol. Biol. (1992) In press
JOURNAL
FEATURES    Location/Qualifiers
            source          1..103
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /cell_line="Ntera2D1"
                        /dev_stage="embryo"
                        /sex="male"
                        /tissue_type="carcinoma"
BASE COUNT  25 a 27 c 33 g 18 t
ORIGIN

Query Match      0.5%; Score 84.4; DB 9; Length 103;
Best Local Similarity 89.2%; Pred. No. 7.3e-05;
Matches 91; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 15007 CTGAGTGCAGTGTGGGATCTCGGCTCAGTCACTGCAACCTCTGCCCTCCGGGTTCAAGTGAT 15066
            |||||||GTTTCACCTTGTGTAGCCAGGAGTGTCGATCTCCGACTCGTGATCC |||||
Db 103 CTGAGTGCAGTGTGGGATCTCGGCTCAGTCACTGCAACCTCTGCCCTCCGGGTTCAAGTGAT 44

Qy 15067 TCTCCTGCCTCAGCTCCCGAGTAGCTGGACTAGCTGGACTACAGGCACA 15108
            |||||||GTTTCACCTTGTGTAGCCAGGAGTGTCGATCTCCGACTCGTGATCC |||||
Db 43 TCTCCTGCCTCAGCTCCCGAGTAGCTGGACTAGCTGGACTACAGGCACA 2

RESULT 4
HSLDLRN2
LOCUS      108 bp      DNA      PRI      20-MAY-1992
DEFINITION Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION  X05250
VERSION     X05250.1 GI:34337
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Homnidae; Homo.
```

REFERENCE	1 (bases 1 to 108)									
AUTHORS	Horschmanke,B., Beisiegel,U., Dunning,A., Havinga,J.R., Williamson,R. and Humphries,S.									
TITLE	Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia									
JOURNAL	Eur. J. Biochem. 164 (1), 77-81 (1987)									
MEDLINE	87161901									
COMMENT	See X05252 for deletion junction									
FEATURES	Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.									
source	Location/Qualifiers									
	1..108									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
intron	1..108									
	/note="intron XIV fragment"									
BASE COUNT	28 a 23 c 39 g 18 t									
ORIGIN										
	Query Match 0.5%; Score 84; DB 10; Length 108;									
	Best Local Similarity 86.1%; Pred. No. 8.4e-05;									
	Matches 93; Conservative 0; Mismatches 15; Indels 0; Gaps 0;									
QY	1765 AAAAAATTAGCTGGCGAGTGCACGACACTCTAATCCAGCTACTCGGGAGGCTGAG 1824									
Db	1 ACMAAAATTAACCCAGCGTGTGGCGAGTGGCTTAATCCAGCTACTCGGGAGGCTGAG 60									
QY	1825 GGAGGACAATCACTGAACCTCAGGAGCAGAGGTTGTAGTCAGCTGAG 1872									
Db	61 GCAGGAGAAATGCTTGAACCCAGGAGGAGGAGTTGCAGTGAGCCGAG 108									
RESULT	5									
HSLDL12	HSLDL12 108 bp DNA PRI 20-MAY-1992									
LOCUS	Human LDL-receptor gene Intron 12 fragment (normal gene) LDL = low density lipoprotein.									
ACCESSION	X05248									
VERSION	X05248.1 GI:34334									
KEYWORDS	Alu repetitive sequence; low density lipoprotein receptor; repetitive sequence.									
SOURCE	human.									
ORGANISM	Homo sapiens									
	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.									
REFERENCE	1 (bases 1 to 108)									
AUTHORS	Horschmanke,B., Beisiegel,U., Dunning,A., Havinga,J.R., Williamson,R. and Humphries,S.									
TITLE	Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia									
JOURNAL	Eur. J. Biochem. 164 (1), 77-81 (1987)									
MEDLINE	87161901									
COMMENT	see X05249 for deletion junction									
FEATURES	Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.									
source	Location/Qualifiers									
	1..108									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
misc_feature	complement(<1..65)									
	/note="Alu repeat"									
intron	1..108									
	/note="intron XII fragment"									
BASE COUNT	21 a 38 c 20 g 29 t									
ORIGIN										
	Query Match 0.5%; Score 83; DB 10; Length 108;									
	Best Local Similarity 86.08; Pred. No. 0.00012;									
	Matches 92; Conservative 0; Mismatches 15; Indels 0; Gaps 0;									
QY	13476 TTGGCTCACTGCAACCTTTGGCTCTTGGGTTCAAGCAATATCTCGCTCAGCCTCCTTA 13535									

Db	2	TCGCCTCACCACACACCTCTGCCTCTGGTTCAAACCATTTTCTGCTCAGCCTCCTTA	61
Qy	13536	GTACAGGGTCTACAGCAGTGTGCCACACACACCGGCTGTTTGTGA	13582
Db	62	GTAGCTGGATTACAGCATGTGCCACACCGCGCTGATTGTGA	108
RESULT	6		
HUMALCE162/c			
LOCUS	HUMALCE162	107 bp ss-RNA	PRI 15-APR-1994
DEFINITION	Human carcinoma cell-derived Alu RNA transcript, clone CE162.		
ACCESSION	M87924		
VERSION	M87924.1	GI:174871	
KEYWORDS	Alu repeat.		
SOURCE	Homo sapiens male embryo carcinoma cDNA to other RNA.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
AUTHORS	Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 107)		
JOURNAL	Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.		
FEATURES	Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences		
source	J. Mol. Biol. (1992) In press		
BASE COUNT	28 a 30 c 35 g 14 t		
ORIGIN	Location/Qualifiers		
	1..107		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/cell_line="Ntera2D1"		
	/dev_stage="embryo"		
	/sex="male"		
	/tissue_type="carcinoma"		
Query Match	0.5%; Score 82.2; DB 9; Length 107;		
Best Local Similarity	87.4%; Pred. No. 0.00017;		
Matches	90; Conservative 0; Mismatches 13; Indels 0; Gaps 0;		
Qy	14973	TTTTTTTGACTGAGTCTGCTCTCTCTCAGAGCTGGAGTGCAGTGGCGATCTCGC	15032
Db	107	TTTTTTGAGCGGAGTCTCGCTCTGTGCCAGGCTGGAGTGCAGTGGCGGATCTCGC	48
Qy	15033	TCACTGCAACCTTCGCTCCCGGGTTTCAAGTGATTCCTCTCC 15075	
Db	47	TCACTGCAAGCTCGCGCTCCCGGGTTCAAGCCATCTCTCTGCC	5
RESULT	7		
HSLDLRD1			
LOCUS	HSLDLRD1	108 bp DNA	PRI 20-MAY-1992
DEFINITION	Human LDL-receptor mutated gene with intron 12 deletion junction.		
ACCESSION	X05249		
VERSION	X05249.1	GI:34335	
KEYWORDS	Alu repetitive sequence; low density lipoprotein receptor.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;		
AUTHORS	Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 108)		
JOURNAL	Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,		
MEDLINE	Williamson,R. and Humphries,S.		
COMMENT	Unequal crossing-over between two alu-repetitive DNA sequences in the low-density lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia		
	Eur. J. Biochem. 164 (1), 77-81 (1987)		
	87161901		
	*source: hypercholesterol aemia		
	See X05248 for corresponding normal gene sequence		
	In the defective LDL-receptor gene the deletion occurred between two alu-repetitive sequences, that are in the same direction, the deletion eliminates exons 13 and 14 and changes the reading frame		

[illegible]

See X05250 for corresponding normal gene sequence
In the defective LDL-receptor gene the deletion occurred between two
alu-repetitive sequences, that are in the same direction, the
deletion eliminates exons 13 and 14 and changes the reading frame
of the resulting spliced mRNA.
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.

```

FEATURES
  source
    1. .108
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /cell_type="blood leukocytes from a patient with familial"
  intron
    1. .108
    /note="intron XIV fragment"
  BASE COUNT      28 a   20 c   40 g   20 t
  ORIGIN

Query Match      0.4%; Score 79.8; DB 10; Length 108;
Best Local Similarity 84.1%; Pred. No. 0.00042;
Matches 90; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 15730 TTGGCTCATTCACACCTCGCTCCTGGTCAAGTCAATTCCTGACTCAGCCTCCCGA 15789
Db 107 TCGCCTCACCACACCTCGCTCCTGGTCAAAACCATTTCTGCTCAGCCTCCCGA 48
Qy 15790 GTAGCTGGGATTACAGGCATGCATCCATGCCCTGGGTAAATTTTGT 15836
Db 47 GTAGCTGGGATTACAGGCATGCCACACGCCCTGGGTAAATTTTGT 1

RESULT 11
HS8IC8R/c      103 bp      DNA      STS      05-SEP-1991
LOCUS          Human sequence tagged site 8IC8R DNA from 19q13.
DEFINITION     X57789
ACCESSION      X57789.1 GI:23938
VERSION         STS; myotonic dystrophy.
KEYWORDS        human.
SOURCE          Homo sapiens
ORGANISM        Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
                Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 103)
AUTHORS        Aldridge,F.L.
TITLE          Direct Submission
JOURNAL        Submitted (12-FEB-1991) F.L. Aldridge, ICI Pharmaceuticals,
                Alderley Park, Macclesfield, Cheshire, SK10 4TG, UK
REFERENCE      2 (bases 1 to 103)
AUTHORS        Butler,R., Riley,J.H., Ogilvie,D.J., Anand,R., Buxton,J.,
                Davies,J., Johnson,K. and Markham,A.F.
TITLE          Two sequence-tagged sites defining the ends of a 380 kb YAC clone
                from 19q13
JOURNAL        Nucleic Acids Res. 19 (17), 4787 (1991)
MEDLINE        91367697
COMMENT        See also X57788 for STS 8IC8L.
FEATURES
  source
    1. .103
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="19q13"
    /germline
    /clone_lib="YAC library: ICI"
    /clone="8IC8"
  BASE COUNT      29 a   28 c   23 g   22 t   1 others
  ORIGIN

Query Match      0.4%; Score 79.2; DB 13; Length 103;
Best Local Similarity 86.1%; Pred. No. 0.00054;
Matches 87; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1556 TACCTGTAATCCCGACACTTTGGGAGACTGAGGTGGGTGATCACTTGAGTCAGGAGTT 1715
Db 102 TGCCTATAATTCGCACTTTGGGAGGTCGAGGTGGGTGGATCACTTAAGGTCAGGAGTT 43

```

```

Qy 1716 CAAACACAGCTGCGCAACATGGTGAACACCACTCTCTACT 1756
Db 42 CTTGACCAGCTGCGCAACATGGTGAACACCACTCTCTACT 2

RESULT 12
HSU67804/c      108 bp      RNA      PRI      01-AUG-1997
LOCUS          Human small cytoplasmic Alu transcript.
DEFINITION     U67804
ACCESSION      U67804.1 GI:2289918
VERSION         U67804.1
KEYWORDS        Alu.
SOURCE          human.
ORGANISM        Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
                Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 108)
AUTHORS        Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE          cDNAs derived from primary and small cytoplasmic Alu (sAlu)
                transcripts
JOURNAL        J. Mol. Biol. 271 (2), 222-234 (1997)
MEDLINE        97415756
REFERENCE      2 (bases 1 to 108)
AUTHORS        Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE          Direct Submission
JOURNAL        Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
                Children's Hospital of Philadelphia, 1004F Abramson Research
                Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES
  source
    1. .108
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="TSCAlu3"
    /note="sAlu"
    /rpt_family="Alu"
    /rpt_type="dispersed"
  repeat_region
    1. .108
  BASE COUNT      26 a   38 c   26 g   18 t
  ORIGIN

Query Match      0.4%; Score 77.8; DB 11; Length 108;
Best Local Similarity 87.6%; Pred. No. 0.00091;
Matches 85; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 15143 GTAGAGATGGAGTTTCGCCGTGTAGCCAGGATGTCGATCTCCTGACCTCGTGATCC 15202
Db 97 GGAAGACGGGGTTTCACCATGTTAGCCAGGATGTCGATCTCCTGACCTTGATCC 38
Qy 15203 ACCGGCTCGCGCTCCCAAGTCTGGGATTACAGGC 15239
Db 37 TCCCGCTTTGGCCTTCCAAAGTCTGGGATTACAGGC 1

RESULT 13
HUMUT8164A/c      91 bp      DNA      STS      29-DEC-1994
LOCUS          Human STS UT8164, 5' primer bind, sequence tagged site.
DEFINITION     L30244
ACCESSION      L30244.1 GI:605447
VERSION         L30244.1
KEYWORDS        STS; PCR primer; STS sequence; microsatellite DNA; microsatellite
                marker; sequence tagged site; tetranucleotide repeat.
SOURCE          Homo sapiens DNA.
ORGANISM        Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
                Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 91)
AUTHORS        Gerken,S.C., Matsunami,N., Plaetke,R., Albertsen,H., Ballard,L.,
                Mellis,F., Lawrence,E., Moore,M., Holik,P.R., Carlson,M., Zhao,X.,
                Robertson,M., Bradley,P., Elsner,T., Tingey,A., Lalouel,J.-M. and
                White,R.
TITLE          Genetic and physical mapping of simple sequence repeat containing

```

JOURNAL COMMENT	sequence tagged sites from the human genome Unpublished (1994) Submitted by: Utah Center for Human Genome Research University of Utah, Dept. of Human Genetics 2160 Eccles Institute of Human Genetics Salt Lake City, UT 84112 e-mail: sts@corona.med.utah.edu Primer A: AGAGGTGCAGTGACACNA Primer B: TTTTCCCCCTCTACTCTACT End to Label: Primer B PCR Profile: Initial Denaturation: 94C 300sec Cycles Denaturation Annealing Extension 5 94 C 10 sec. 56 C 10 sec. 72 C 20 sec. 30 94 C 10 sec. 52 C 10 sec. 72 C 20 sec. Mg++: 1.50 mM Gel: Acrylamide 7%, Formamide 32%, Urea 34%									
FEATURES	Alleles: 1. Location/Qualifiers 1. .91 /organism="Homo sapiens" /db_xref="taxon:9606" 10. .28 /evidence=experimental 35 a 20 c 23 g 13 t									
BASE COUNT	35 a 20 c 23 g 13 t									
ORIGIN										
Query Match	0.4%; Score 77.2; DB 13; Length 91;									
Best Local Similarity	91.1%; Pred. No. 0.0012;									
Matches	82; Conservative 0; Mismatches 8; Indels 0; Gaps 0;									
Qy 14965	TTTTTTTTTTTTTTTCTGACTGAGCTTGCTCTGCTGTCTACAGCTGGAGTGAGTGGTGCG 15024 									
Db 91	TTTTTTTTTTTTTTTGTACAGAGTCTACTCTGTCAACCCAGCTGGAGTGAGTGGTG 32 									
Qy 15025	ATCTCGGCTCACTGCAACCTCTGCCTCCG 15054 									
Db 31	ATCTTGGTTCACTGCAACCTCTGCCTCCG 2 									
RESULT 14										
HSU67807	HSU67807 110 bp RNA PRI 01-AUG-1997									
LOCUS	Human small cytoplasmic Alu transcript.									
DEFINITION	U67807									
ACCESSION	U67807.1 GI:2389921									
VERSION	U67807.1									
KEYWORDS	Alu.									
SOURCE	human.									
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE	1 (bases 1 to 110) Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L. cDNAs derived from primary and small cytoplasmic Alu (scAlu) transcripts									
AUTHORS	J. Mol. Biol. 271 (2), 222-234 (1997)									
TITLE	JOURNAL									
JOURNAL	MEDLINE 97415756									
REFERENCE	2 (bases 1 to 110) Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L. Direct Submission									
AUTHORS	Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The Children's Hospital of Philadelphia, 1004F Abramson Research Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA									
TITLE	Location/Qualifiers 1. .110 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="TscAlu6" 1. .110 /note="scAlu" /rpt_family="Alu" /rpt_type=dispersed 26 a 39 c 24 g 21 t									
FEATURES										
source										
repeat_region										
BASE COUNT	26 a 39 c 24 g 21 t									

Search completed: June 20, 2000, 17:57:22
Job time: 537526 sec

CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.

Query Match 0.4%; Score 71.8; DB 1; Length 108;
Best Local Similarity 84.9%; Pred. No. 0.019;
Matches 79; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

Qy 15834 TGTATTTTGTAGAGATGGGTTTCCACCATGTGACAGGCTGTCTCAAACTCTGTGAC 15893
Dd 1 TGTCTTTTGTAGAGATGGGTTTCTCTGTGTCGAGGATGTCTCGAACTCTGTGAC 60

Qy 15894 CTCAGTGCATCCGCTGCTAGCTCCCAAAA 15926
Dd 61 TTCAGTGTATCCGCTGCTGCTGCTCCCAAAA 93

RESULT 3
X12095/C
ID X12095 standard; DNA; 108 BP.
AC X12095;
DE Human biallelic polymorphic DNA fragment TIGR-A003M18a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease

CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, Ehlers-Danlos
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.

Query Match 0.4%; Score 68.6; DB 1; Length 108;
Best Local Similarity 83.8%; Pred. No. 0.055;
Matches 88; Conservative 1; Mismatches 15; Indels 1; Gaps 1;

Qy 6421 TAATCCCGACAC-TTTGGAGGCGGAGGAGGATCATCATGAGTCAAGAAATTTCAAGA 6479
Dd 105 TAATCCCGACAC-TTTGGAGGCGGAGGAGGATCATCATGAGTCAAGAAATTTCAAGA 46

Qy 6480 CCAGCTGGCCAAATGGTGAACCTCATCTCTACTAAAAATACA 6524
Dd 45 CCATCTCTGGCCAAACAYAGGAAACCTCATCTCTACTAAAAAAGACA 1

RESULT 4
X124892/C
ID X124892 standard; cDNA to mRNA; 100 BP.
AC X124892;
DE Human gene signature HUMGS06998.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATSU) MATSUBARA K.
PI Matsubara K, Okubo K.
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 1720; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

SQ Sequence 100 BP; 28 A; 23 C; 28 G; 37 T;

CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.

Query Match 0.4%; Score 68.6; DB 1; Length 108;
Best Local Similarity 83.8%; Pred. No. 0.055;
Matches 88; Conservative 1; Mismatches 15; Indels 1; Gaps 1;

Qy 6421 TAATCCCGACAC-TTTGGAGGCGGAGGAGGATCATCATGAGTCAAGAAATTTCAAGA 6479
Dd 105 TAATCCCGACAC-TTTGGAGGCGGAGGAGGATCATCATGAGTCAAGAAATTTCAAGA 46

Qy 6480 CCAGCTGGCCAAATGGTGAACCTCATCTCTACTAAAAATACA 6524
Dd 45 CCATCTCTGGCCAAACAYAGGAAACCTCATCTCTACTAAAAAAGACA 1

RESULT 4
X124892/C
ID X124892 standard; cDNA to mRNA; 100 BP.
AC X124892;
DE Human gene signature HUMGS06998.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATSU) MATSUBARA K.
PI Matsubara K, Okubo K.
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 1720; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

SQ Sequence 100 BP; 28 A; 23 C; 28 G; 37 T;

Query Match	0.4%	Score 65.6;	DB 1;	Length 100;	
Best Local Similarity	77.8%	Pred. No. 0.14;	Mismatches 0;	Indels 0;	Gaps 0;
Matches	77;	Conservative			

QY	14968	TTTTTTTTTTTTTTGACATGAGTCTTGCTCTCACAGGCTGGAGTGCAGTGGTGCATC	15027
DB	100	TTTTTTTGTTCAAACAGAGTGTCACTCTGTCAACCCAGGCGAGTGCANGGTGCAATC	41
QY	15028	TGCGCTCACTGACACCTCTGCCTCCCGGTTCAAGTGAT	15066
DB	40	TCAGCTNATGTGCAAAATCTGCCTCCAGGTTCAAGCGAT	2

RESULT	5
T26213/c	
ID	T26213 standard; cDNA to mRNA; 103 BP.
AC	T26213;
DT	13-NOV-1996 (first entry)
DE	Human gene signature HUMGS08452.
KW	Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW	human; cloning; mapping; non-biased library; diagnosis; detection;
KW	cell typing; abnormal cell function; ss.
OS	Homo sapiens.
PN	W09514/72-A1.
PD	01-JUN-1995.
PF	11-NOV-1994; J01916.
PR	12-NOV-1993; JP-355504.
PA	(MATS/) MATSUBARA K.
PA	(OKUB/) OKUBO K.
PI	Matsubara K, Okubo K;
PT	WPI; 95-206931/27.
DR	Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT	for diagnosis of abnormal cell function, by preparing cDNA that
PT	reflects relative abundance of corresp. mRNA in specific human
PT	tissues
PS	Claim 1; Page 2029; 2245pp; Japanese.
CC	A single-stranded DNA (or its complementary strand or the corresp.
CC	double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC	given in T19001-T26837 and which is able to hybridise to part of
CC	human genomic DNA, cDNA or mRNA is claimed. The GS (gene Signature)
CC	sequences were obtained from 3'-directed cDNA libraries prepared
CC	from various human tissues; synthesis of cDNA was initiated from the
CC	3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC	untranslated sequence is unique to a particular mRNA species, almost
CC	all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC	is constructed so as to reflect accurately the relative abundance of
CC	different mRNAs in the particular tissue from which it was derived.
CC	The appearance frequency of a given GS in a cDNA library can be
CC	determined (esp. using primers and probes derived from the GS
CC	sequences) as a means of diagnosing abnormal cell function or for
CC	recognising different cell types.
SQ	Sequence 103 BP; 33 A; 25 G; 23 T;

Query Match	0.4%	Score 65.2;	DB 1;	Length 103;	
Best Local Similarity	77.5%	Pred. No. 0.16;	Mismatches 0;	Indels 0;	Gaps 0;
Matches	79;	Conservative			

QY	8118	TTTCTTTTTTCTTGACAGGCTCTTGCTCTATTGCTAGGCTGAGTGCAGTGGTGCAC	8177
DB	102	TTTTTTTTTCTTAAAGACATGTTCTTACTCTGTGCCCCAGGCTGGAGTGCAGTGGTGCA	43
QY	8178	TCATGCTCACTGAGCGCTTGAACCTCCAGGCTCAAGCAATC	8219
DB	42	TCATAGTCACTGTGAACCAACCAACTCTGGACTCAAGTGATC	1

RESULT	6
T66081	
ID	T66081 standard; DNA; 92 BP.
AC	T66081;

PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1: Page 1720: 2245pp: Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;

Query Match 0.4%; Score 64; DB 1; Length 100;
Best Local Similarity 76.8%; Pred. No. 0.24;
Matches 76; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1833 ATCACTTGAACTCAGGAGGAGTGTAGTGAGCTGAGATCGACCTGCACCTCCAG 1892

Db 2 ATCCCTTGAACTCGGAGGAGAAATTCGCAATNAGCTGAGATTGCACCTTGCACTCCG 61

QY 1893 CCTGGGTGACAGCAAGACTCCATTTTAAAAAATAATA 1931

Db 62 CCTGGGTGACAGAGTGACACTCTCTTTTGAACAAACAA 100

RESULT 8

T25009/c
ID T25009 standard; cDNA to mRNA; 108 BP.

AC T25009;
DT 07-NOV-1996 (first entry)
DE Human gene signature HUMGS07131.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K. Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1: Page 1748: 2245pp: Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

SQ Sequence 108 BP; 34 A; 31 C; 26 G; 15 T;

Query Match 0.4%; Score 64; DB 1; Length 108;
Best Local Similarity 74.5%; Pred. No. 0.24;
Matches 79; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 17285 TTGTGTTGTTGTTTGTGATAGAGTCTTCTCTCATTCAGGTGAGTGCAGTGG 17344

Db 107 TTGNTGTGTTGTTGTTTCAACAGGCTCTGTCTCTCACTCAGGTGGATNACATGG 48

QY 17345 CATGATCTCAGCTACCTGCAGCCTCCGCCCTCCCGGGTTCAAGAGAT 17390

Db 47 CGTGACCATGGCTCACTGCAGCCTTGCCCTCATGGGCTCAGGGCAT 2

RESULT 9

T26828
ID T26828 standard; cDNA to mRNA; 108 BP.

AC T26828;
DT 14-NOV-1996 (first entry)
DE Human gene signature HUMGS09078.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K. Okubo K;
DR WPI; 95-206931/27.

PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1: Page 2182: 2245pp: Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

SQ Sequence 108 BP; 18 A; 33 C; 23 G; 28 T;

Query Match 0.4%; Score 63.8; DB 1; Length 108;
Best Local Similarity 88.3%; Pred. No. 0.26;
Matches 68; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 15182 GATCTCTGACCTCGTGATCCACCGCTCGGCTCCCAAGTCTGGATTCAGGCAT 15241

Db 1 GATCTCTGACCTCGTGATCCCGCCCTGCTGCCATAGTGTGGGNTTACAGGCAT 60

QY 15242 GGGCCACACGCTGGC 15258

Db 61 GAGCCACACGCGCGG 77

RESULT 10

V11611/c
ID V11611 standard; cDNA; 106 BP.

CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 87 BP; 35 A; 21 C; 16 G; 13 T

Query Match 0.3%; Score 58.4; DB 1; Length 87;
Best Local Similarity 79.1%; Pred. No. 1.5;
Matches 68; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 15819 TGCCTGGGTAATTTTGTATTTTAGTAGAGATGGGGTTTCACCATGTTGACCAAGGCTGG 15878

Db 86 TTCTTGGCTNATTTCTGTATTTTTGTAAGATGGGTTTCGCCATGTTCTCTCGGGCTGG 27

QY 15879 TCTCAAACTCCTGACCTCAAGTGATC 15904

Dbb 26 TTTTAACTCCTGGGNTCAAGCGATC 1

Search completed: June 20, 2000, 18:25:46
Job time: 538057 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2000, 09:51:12 ; Search time 13789.4 Seconds
(without alignments)
5291.175 Million cell updates/sec

Title: US-08-852-495C-2_COPY_196000_214000
Perfect score: 18001
Sequence: 1 CTRAAAAGTATTTTAAACC.....TGGACATGCTGTGTTCTTC 18001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues
Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result % Query
SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	93	0.5	109	30	AA243009	AA243009 zr25h02.s
2	91.6	0.5	106	37	AA703692	AA703692 ac81a10.r
3	91	0.5	107	35	AA585533	AA585533 nk42b11.s
c 4	90.2	0.5	103	84	B48914	B48914 RPC111-4A12
c 5	90	0.5	110	106	AQ386882	AQ386882 RPC111-13
c 6	88.6	0.5	103	108	AQ335244	AQ335244 RPC11-13
c 7	87.6	0.5	110	106	AQ386882	AQ386882 RPC11-13
c 8	87.4	0.5	106	108	AQ544957	AQ544957 CITBT-EI-
c 9	87.2	0.5	109	84	B17434	B17434 345K2.rvb C
c 10	86.8	0.5	110	38	AA897366	AA897366 am06h02.s
c 11	86.6	0.5	103	38	AA807640	AA807640 nx08h05.s
c 12	86.6	0.5	110	30	AA244245	AA244245 nc07a04.s
c 13	86.6	0.5	110	94	AQ003188	AQ003188 RPC111-1D
c 14	86.2	0.5	107	35	AA565533	AA565533 nk42b11.s
c 15	86.2	0.5	108	84	B65160	B65160 CIT-HSP-201
c 16	86	0.5	105	109	AQ637292	AQ637292 RPC11-14
c 17	85.4	0.5	103	94	AQ028649	AQ028649 CIT-HSP-2
c 18	85.4	0.5	103	108	AQ335244	AQ335244 RPC11-13
c 19	85	0.5	106	30	AA250812	AA250812 zs06a05.s
c 20	85	0.5	109	84	B17434	B17434 345K2.rvb C
c 21	84.4	0.5	102	36	AA654562	AA654562 nt75f10.s
c 22	83.6	0.5	106	105	AQ264176	AQ264176 CITBT-EI-
c 23	83.4	0.5	109	105	AQ265749	AQ265749 CITBT-EI-
c 24	83	0.5	107	35	AA583252	AA583252 nk41e04.s
c 25	83	0.5	109	94	AQ029690	AQ029690 RPC111-41
c 26	82	0.5	103	108	AQ582186	AQ582186 RPC11-14
c 27	82	0.5	106	63	AI991750	AI991750 wt48e01.x
c 28	82	0.5	106	63	AI991750	AI991750 wt48e01.x
c 29	81.6	0.5	107	24	H67040	H67040 yu68c01.r1
c 30	81.2	0.5	102	84	B48088	B48088 RPC111-4N6.
c 31	81.2	0.5	104	108	AQ544583	AQ544583 CITBT-EI-
c 32	81.2	0.5	108	35	AA594869	AA594869 nc021e02.s
c 33	81	0.4	105	28	AA078003	AA078003 7H12D08.C
c 34	81	0.4	107	103	AQ240182	AQ240182 CIT-HSP-2
c 35	80.8	0.4	108	84	B32951	B32951 HS-1016-A1
c 36	80.8	0.4	109	103	AQ200347	AQ200347 RPC111-43
c 37	80.6	0.4	103	35	AA570438	AA570438 nk63g02.s
c 38	80.6	0.4	103	108	AQ534922	AQ534922 RPC11-13
c 39	80.6	0.4	108	84	B15423	B15423 345B10.TV C
c 40	80.4	0.4	106	38	AA812141	AA812141 OB48h02.s
c 41	80.4	0.4	109	24	N25299	N25299 yw52c09.s1
c 42	80.2	0.4	106	30	AA250812	AA250812 zs06a05.s
c 43	80.2	0.4	109	94	AQ028426	AQ028426 CIT-HSP-2
c 44	80.2	0.4	109	105	AQ265749	AQ265749 CITBT-EI-
c 45	80	0.4	97	39	AA837701	AA837701 oe06c02.s

ALIGNMENTS

```

RESULT 1
AA243009      AA243009      109 bp      mRNA      EST      11-MAR-1998
LOCUS         zr25h02.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
DEFINITION    cDNA clone IMAGE:664467 3' similar to contains Alu repetitive
               element;contains element LTR1 repetitive element ;, mRNA sequence.
ACCESSION     AA243009
VERSION       AA243009.1 GI:1873869
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
               Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;
               Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 109)
AUTHORS       Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
               Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
               Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
               Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
               WashU-NCI human EST Project
TITLE         Unpublished (1997)
JOURNAL
COMMENT       On Dec 3, 1996 this sequence version replaced gi:1126869.

```

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1127 Std Error: 0.00
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 102.

FEATURES

Location/Qualifiers
1..109
/organism="Homo sapiens"
/db_xref="GDB:5426481"
/db_xref="taxon:9606"
/clone="IMAGE:664467"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/tissue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: brain; Vector: pBluescript SK-; Site:1;
ECORI; Site2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Uninduced, exponentially growing neuroepithelial
cells (Ntera-2/ci.D1). Average insert size: 1.0 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

BASE COUNT 19 a 30 c 30 g 30 t

ORIGIN

Query Match 0.5%; Score 93; DB 30; Length 109;
Best Local Similarity 90.8%; Pred. No. 0.066;
Matches 99; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Oy 15134 GTATTTTGTAGATGAGATGATTCGCGCTGTAGCCAGGATGCTCGATCTCCTGACC 15193
|||||
Db 1 GTATTTTGTAGATGAGATGATTCACCGTGTAGCCAGGATGCTCGATCTCCTGACC 60
Oy 15194 TCGTGATCCACGGCTCGGCTCCCAAGTCTCGGATTCACGCGCATG 15242
|||||
Db 61 TCGTGATCCGCGCCACTCGGCTCCCAAGTCTCGGATTCACGCGCATG 109

RESULT 2

AA703692 106 bp mRNA EST 24-DEC-1997
LOCUS aa81a10.r1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone
DEFINITION IMAGE:1140858 5' similar to contains Alu repetitive element; , mRNA
sequence.
ACCESSION AA703692
VERSION AA703692.1 GI:2713610
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 106)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
 Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
 Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
 WashU-NCI human EST Project
TITLE Unpublished (1997)
JOURNAL
COMMENT On Sep 12, 1996 this sequence version replaced gi:1397630.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 53.

FEATURES

source

1. 106 Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1140858"
/clone_lib="Stratagene hnt neuron (#937233)"
/dev_stage="hnt neurons"
/lab_host="SOLR (kanamycin resistant)"
/note="vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Differentiated, post mitotic hnt neurons. Average insert size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTGTGTTTTTTTTT 3'."
CTCGAGTGTGTTTTTTTTT 3'."

19 a 29 c 29 g 29 t

BASE COUNT

ORIGIN

Query Match 0.5%; Score 91.6; DB 37; Length 106;
Best Local Similarity 91.5%; Pred. No. 0.097; Mismatches 0; Indels 0; Gaps 0;
Matches 97; Conservative 0;

QY 15137 TTTTGTAGATGAGGTTTCGCGCTGTAGCCAGGATGCTCGATCTCCTGACCTCG 15196

Db 1 TTTTGTAGATGAGGATGTTTACCGTGTAGCCAGGATGCTCGATCTCCTGACCTCG 60

QY 15197 TGATCAGCGGCTCGGCTCCCAAGTGTGGGATTACAGGCATG 15242

Db 61 TGATGCGCGCGCTCAGCTCCCAAGTGTGGGATTACAGGCATG 106

RESULT 3

AA565533

LOCUS

AA565533 107 bp mRNA EST 08-SEP-1997
n42b11.s1 NCI-CCAP_GC2 Homo sapiens cDNA clone IMAGE:1016157 3'
similar to contains Alu repetitive element;; mRNA sequence.

ACCESSION

VERSION

AA565533.1

GI:2337172

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

NCI-CCAP

http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

On Sep 12, 1996 this sequence version replaced gi:1393355.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc., David B. Krizman, Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
Clone Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 1661 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 87.

FEATURES

source

1. 107 Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1016157"
/clone_lib="NCI-CCAP_GC2"

/tissue_type="germ cell tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Bulk germ cell tumor. 5' adaptor sequence: 5' GAATTCGCGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTGTGTTTTTTTTT 3'."
Average insert size: 1.2 kb."

22 a 34 c 26 g 25 t

BASE COUNT

ORIGIN

Query Match 0.5%; Score 91; DB 35; Length 107;
Best Local Similarity 90.7%; Pred. No. 0.11; Mismatches 97; Conservative 0; Indels 0; Gaps 0;

QY 15007 CTGGAGTCAGTGGTCCGATCTCGGCTCACTGCAACCTCTGCTCCCGGTTCAAGTGAT 15066

Db 1 CTGGAGTCAGTGGTCCGATCTCAATCTCAGCTCACTGCAACCTCTGCTCCCGGTTCAAGTGAT 60

QY 15067 TCTCTGCTCAGCTCCGAGTAGCTTGGACTACAGGCACACCA 15113

Db 61 TCTGCTCAGCTCCGAGTAGCTTGGAGTTACAGGCACACCA 107

RESULT 4

B48914/c

LOCUS

B48914

DEFINITION

RPC111-4A12.TP

103 bp

DNA

GSS

08-APR-1999

ACCESSION

B48914

VERSION

B48914.1

GI:2601151

KEYWORDS

GSS.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and Venter, J.C.

TITLE

Use of BAC End Sequences for Sequence-Ready Map Building

JOURNAL

Unpublished (1997)

COMMENT

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are derived from the human BAC library RPC111. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1. 103

/organism="Homo sapiens"

/db_xref="GDB:7501163"

/db_xref="taxon:9606"

/clone="RPC111-4A12"

/clone_lib="RPC111"

/sex="Male"

/cell_type="Lymphocytes"

/note="vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI; RPC111 Human Male BAC Library"

BASE COUNT

ORIGIN

30 a 28 c 30 g 15 t

Query Match 0.5%; Score 90.2; DB 84; Length 103;

```
Best Local Similarity 92.2%; Pred. No. 0.14;
Matches 95; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 15133 TGTATTTTGTAGAGAGAGTTTCGCCGTTAGCCAGGATGCTCTCGATCTCCTGAC 15192
|||||
Db 103 TGTATTTTGTAGAGAGCGGGTTTACCGTTTATGCGGGATGCTCTCGATCTCCTGAC 44
|||||

Qy 15193 CTCGTGATCCACCGGCTCGGCTCCCAAGTGTGGGATTAC 15235
|||||
Db 43 CTCGTGATCCCGCGCTCGGCTCCCAAGTGTGGGCTTAC 1
|||||

RESULT 5
AQ386882/c
LOCUS
DEFINITION RPCI11-13414_TV RPCI-11 Homo sapiens genomic clone RPCI-11-13414,
genomic survey sequence.
ACCESSION AQ386882
VERSION AQ386882.1 GI:4357905
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Other_GSSs: RPCI11-13414.TJ
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..110
/organism="Homo sapiens"
/db_xref="GDB:7551267"
/db_xref="taxon:9606"
/clone="RPCI-11-13414"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT 26 a 26 c 38 g 20 t
ORIGIN
Query Match 0.5%; Score 90; DB 106; Length 110;
Best Local Similarity 90.6%; Pred. No. 0.14;
Matches 96; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 15853 GGGTTTACCATTGACGAGCTGGTCTCAACTCTCGACTCAAGTATCCACCTGCC 15912
|||||
Db 110 GGGTTTACCATTGTTGTCCAGGCTGGTCTTGAACCTCTGACCTCAAGCGATCCACCTGCC 51
|||||

Qy 15913 TTAGCCTCCCAAAATGCTGGGACTACAGCGGTGAGCCACTGCACC 15958
|||||
Db 50 TCAGCCTCCCAAAAGTACTTGGATTACGGCTGAGCCACTGCTCCC 5
|||||
```

```
RESULT 6
AQ35244
LOCUS
DEFINITION RPCI-11-317H22_TV RPCI-11 Homo sapiens genomic clone
RPCI-11-317H22, genomic survey sequence.
ACCESSION AQ35244
VERSION AQ35244.1 GI:4846934
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..103
/organism="Homo sapiens"
/db_xref="GDB:7621533"
/db_xref="taxon:9606"
/clone="RPCI-11-317H22"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT 31 a 27 c 27 g 18 t
ORIGIN
Query Match 0.5%; Score 88.6; DB 108; Length 103;
Best Local Similarity 91.3%; Pred. No. 0.22;
Matches 94; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 10405 CCAGCATTTCGAGGCTGAGCGGGGAGATCATTGAGGTCAGGATTCGAGACCACC 10464
|||||
Db 1 CCAGCATTTCGAGGCTGAGCGGCAAGCGGAGATCATTGAGGTCAGGATTCGAGACCACC 60
|||||

Qy 10465 TGGCCAAACATGTTGAAACCCCTGCTCCACTAAAAATACAAAA 10507
|||||
Db 61 TGGCCAAACATGTTGAAACCCCGCTCTCTCTATTAATACAAAA 103
|||||

RESULT 7
AQ386882
LOCUS
DEFINITION RPCI11-13414_TV RPCI-11 Homo sapiens genomic clone RPCI-11-13414,
genomic survey sequence.
ACCESSION AQ386882
VERSION AQ386882.1 GI:4357905
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
```

REFERENCE 8
 AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.
 TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
 JOURNAL Unpublished (1997)
 COMMENT Other_GSSs: RPCI11-13414.TJ
 Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbeetigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
 Seq primer: T7
 Class: BAC ends.

FEATURES

source
 Location/Qualifiers
 1..110
 /organism="Homo sapiens"
 /db_xref="GDB:7551267"
 /db_xref="taxon:9606"
 /clone="RPCI-11-13414"
 /clone_lib="RPCI-11"
 /sex="Male"
 /cell_type="Lymphocytes"
 /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"
 BASE COUNT 26 a 26 c 38 g 20 t
 ORIGIN

Query Match 0.5%; Score 87.6; DB 106; Length 110;
 Best Local Similarity 87.3%; Pred. No. 0.27; Mismatches 0; Gaps 0;
 Matches 96; Conservative 0; Indels 0; Gaps 0;
 QY 10375 GGCCGGCGCTGGTGCAGCTGTAATCCAGCACTTTGGCAGGCTGAGCGGCAGCA 10434
 Db 1 GGCCGGAGCAGTGGCTACGCCCTGTAATCCAGTACTTTGGAGGCTGAGCAGGTGGA 60
 QY 10435 TCACTTGAGTCAGGAGTTTGAGACAGCCTGGCCACATGTTGAACCC 10484
 Db 61 TCGCTTGAGTCAGGAGTTTCAGACACCGCTGGACACATGTTGAACCC 110

RESULT 8
 AQ544957/c 106 bp DNA GSS 28-MAY-1999
 LOCUS CITBI-EI-2629N2.TF CITBI-EI Homo sapiens genomic clone 2629N2, genomic survey sequence.
 ACCESSION AQ544957
 VERSION AQ544957.1 GI:4903683
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 106)
 AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., Shizuya, H., Simon, M. and Venter, J.C.
 TITLE Use of BAC End Sequences from Caltech Libraries for Sequence-Ready Map Building
 JOURNAL Unpublished (1997)
 COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbeetigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html.
 Seq primer: M13-21
 Class: BAC ends.

FEATURES

source
 Location/Qualifiers
 1..106
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="2629N2"
 /clone_lib="CITBI-EI"
 /sex="male"
 /cell_type="sperm"
 /note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI; CalTech Human BAC Library D"
 BASE COUNT 26 a 26 c 36 g 18 t
 ORIGIN

Query Match 0.5%; Score 87.4; DB 108; Length 106;
 Best Local Similarity 89.5%; Pred. No. 0.29; Mismatches 94; Conservative 0; Indels 0; Gaps 0;
 QY 15155 TTTCGCGGTGTAGCCAGGATGCTCGATCTCTGACCTCGTATCCACCGGCTCGGC 15214
 Db 106 TTTCACCGTGTAGCCAGGATGCTGTGATCTCTGACCTAAATGATCTGCCGCTCGGC 47
 QY 15215 CTCCCAAAGTCTGGGATTACAGGCATGGCCACCGCTGGCC 15259
 Db 46 CTCCCAAAGTCTGGGATTACAGGCTGGGCTCCAGCCCGCGCC 2

RESULT

9
 BI7434 109 bp DNA GSS 04-JUN-1998
 LOCUS 345K2.TVB CIT978SK1 Homo sapiens genomic clone A-345K02, genomic survey sequence.
 ACCESSION BI7434
 VERSION BI7434.1 GI:2125183
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 109)
 AUTHORS Adams, M.D., Kelley, J.M., Rounsley, S.R. and Venter, J.C.
 TITLE Use of a BAC End Sequence Database for Sequence-Ready Map Building
 JOURNAL Unpublished (1997)
 COMMENT Other_GSSs: 345K02.TP 345K02.TPB
 Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
 Seq primer: T7
 Class: BAC ends.

FEATURES

source
 Location/Qualifiers
 1..109
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="A-345K02"
 /clone_lib="CIT978SK1"
 /sex="Female"
 /cell_type="Fibroblast"
 /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;

```

BASE COUNT      24 a      30 c      31 g      24 t
ORIGIN

Query Match      0.5%; Score 87.2; DB 84; Length 109;
Best Local Similarity 88.0%; Pred. No. 0.3;
Matches 95; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1650 GGCTCATACCTGTATATCCAGCAGCTTTGGGAGACTGAGGTGGTGATCATGTGAGGTCA 1709
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 GGCTCATACCTATATATCTAGACATTTGGGAGCTGATGTGGCGGATCAGCTGAGGTGCG 61
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1710 GGAGTTAAAACAGCCTGGCCACATGTTGAAAACCCATCTCTACTA 1757
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 GGAGTTGAGAGCAGCCTGGCCACCATGTTGAAACCCGCTCTCACTA 109
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
AA897366      110 bp      mRNA      EST      04-JAN-1999
LOCUS
DEFINITION   am06h02.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:1466067 3' similar to contains Alu repetitive element;; mRNA
sequence.
ACCESSION   AA897366
VERSION     AA897366.1 GI:3033986
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 110)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     On Jan 19, 1998 this sequence version replaced gi:2150764.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

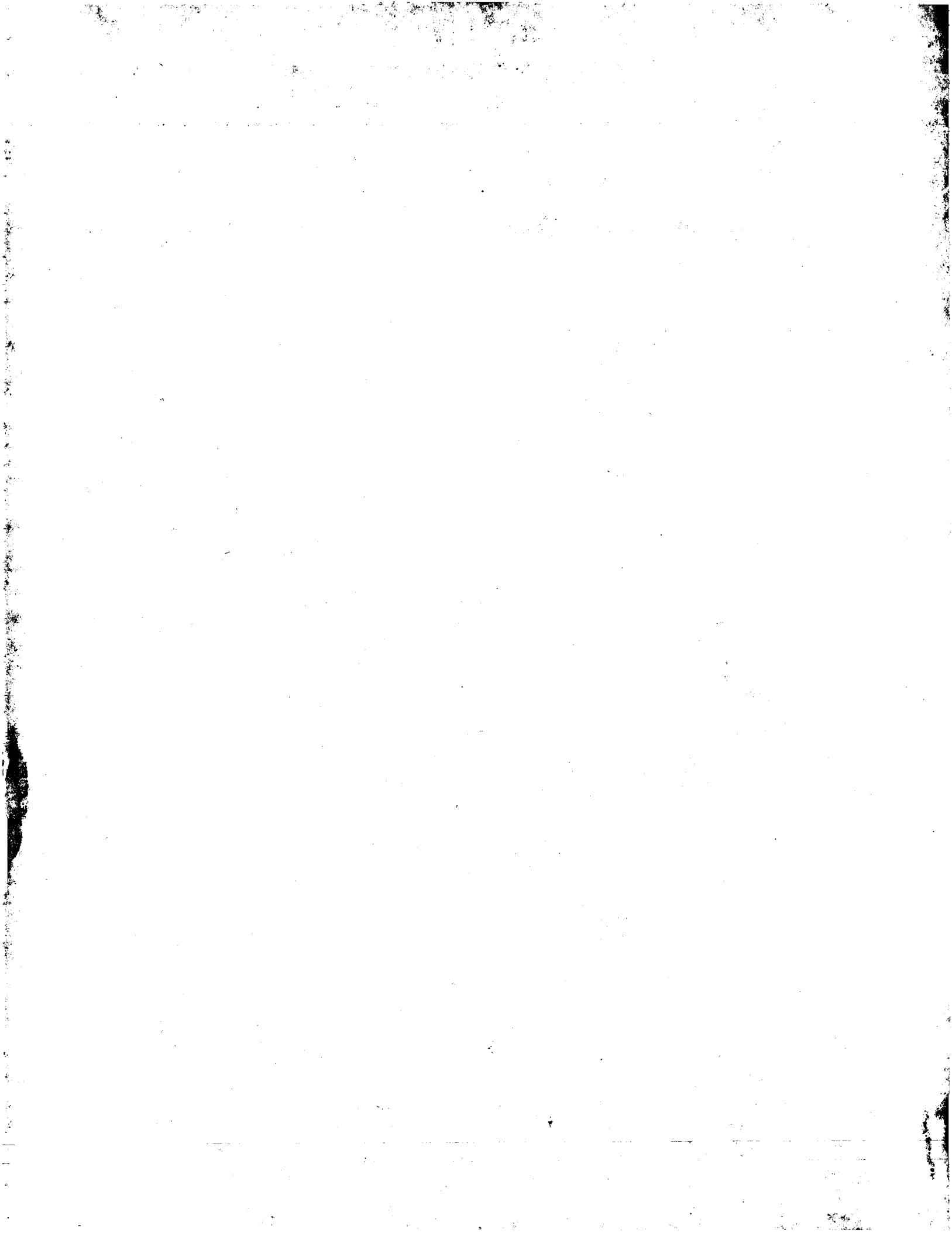
FEATURES             source
Location/Qualifiers
1..110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1466067"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI-CGAP-GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."
22 a      27 c      29 g      32 t

BASE COUNT      22 a      27 c      29 g      32 t
ORIGIN

Query Match      0.5%; Score 86.8; DB 39; Length 110;
Best Local Similarity 88.7%; Pred. No. 0.34;
Matches 94; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 15137 TTTTGTAGATAGGAGTTTCGGCGTGTAGCCAGGATGCTCGATCTCCTGACCTCG 15196
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2000, 18:07:04 ; Search time 599.42 Seconds
(without alignments)
3903.542 Million cell updates/sec

Title: US-08-852-495C-2_COPY_196000_214000
Perfect score: 18001
Sequence: 1 CTAAGAAAGTATTTAAAGC.....TGGACATGCTGTCTTCCTTC 18001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78.4	0.4	105	4	US-08-481-658B-65
2	78.4	0.4	105	4	US-08-477-504A-65
3	78.4	0.4	105	4	US-08-486-756A-65
4	78.4	0.4	105	4	US-08-485-862B-65
5	78.4	0.4	105	5	US-08-787-739-65
6	72.8	0.4	105	4	US-08-481-658B-65
7	72.8	0.4	105	4	US-08-477-504A-65
8	72.8	0.4	105	4	US-08-486-756A-65
9	72.8	0.4	105	4	US-08-485-862B-65
10	72.8	0.4	105	5	US-08-787-739-65
11	64	0.4	92	1	US-08-222-177A-430
12	63.8	0.4	84	3	US-08-454-557C-91
13	63.8	0.4	84	4	US-08-340-426D-91
14	63.8	0.4	84	4	US-08-450-673C-91
15	63.8	0.4	84	6	PCT-US95-17111A-91
16	60.4	0.3	78	3	US-08-454-557C-70
17	60.4	0.3	78	4	US-08-340-426D-70
18	60.4	0.3	78	4	US-08-450-673C-70
19	60.4	0.3	78	6	PCT-US95-17111A-70
20	60.2	0.3	98	1	US-08-088-658-42
21	60.2	0.3	98	4	US-08-471-907A-42
22	59.2	0.3	85	3	US-08-454-557C-92
23	59.2	0.3	85	4	US-08-340-426D-92
24	59.2	0.3	85	4	US-08-450-673C-92
25	59.2	0.3	85	6	PCT-US95-17111A-92
26	59	0.3	60	1	US-08-222-177A-244
27	58.2	0.3	84	3	US-08-454-557C-91

Sequence 91, Appl
Sequence 91, Appl
Sequence 91, Appl
Sequence 70, Appl
Sequence 70, Appl
Sequence 70, Appl
Sequence 44, Appl
Sequence 44, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 166, App
Sequence 57, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-65
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.4%; Score 78.4; DB 4; Length 105;
Best Local Similarity 84.6%; Pred. No. 6.9e-09;

Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 15129 TTTTGTATTTTAGTAGAGATGAGTTTCGCCGTGTTAGCCAGATGGTCTCGATCTCC 15188

Db 2 TTTTACATCTTTAGTAGACAGAGGGTTTACCATATTTGCCAGGCTGCTCTCAAACTCC 61

Qy 15189 TGACCTCGTGATCCACCGGCTCGGCCCTCCCAAAAGTGTGGAT 15232

Db 62 TGACCTGTGTATCCACCAAGCTCGGCCCTCCCAAAAGTGTGGAT 105

RESULT 2

US-08-477-504A-65

; Sequence 65, Application US/08477504A

; Patent No. 5972353

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,504A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3D

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-477-504A-65

Query Match 0.4%; Score 78.4; DB 4; Length 105;

Best Local Similarity 84.6%; Pred. No. 6.9e-09;

Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 15129 TTTTGTATTTTAGTAGAGATGAGTTTCGCCGTGTTAGCCAGATGGTCTCGATCTCC 15188

Db 2 TTTTACATCTTTAGTAGACAGAGGGTTTACCATATTTGCCAGGCTGCTCTCAAACTCC 61

Qy 15189 TGACCTCGTGATCCACCGGCTCGGCCCTCCCAAAAGTGTGGAT 15232

Db 62 TGACCTGTGTATCCACCAAGCTCGGCCCTCCCAAAAGTGTGGAT 105

RESULT 3

US-08-486-756A-65

; Sequence 65, Application US/08486756A

; Patent No. 5981711

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,756A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-486-756A-65

Query Match 0.4%; Score 78.4; DB 4; Length 105;

Best Local Similarity 84.6%; Pred. No. 6.9e-09;

Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 15129 TTTTGTATTTTAGTAGAGATGAGTTTCGCCGTGTTAGCCAGATGGTCTCGATCTCC 15188

Db 2 TTTTACATCTTTAGTAGACAGAGGGTTTACCATATTTGCCAGGCTGCTCTCAAACTCC 61

Qy 15189 TGACCTCGTGATCCACCGGCTCGGCCCTCCCAAAAGTGTGGAT 15232

Db 62 TGACCTGTGTATCCACCAAGCTCGGCCCTCCCAAAAGTGTGGAT 105

RESULT 4

US-08-485-862B-65

; Sequence 65, Application US/08485862B

; Patent No. 5989838

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,862B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-485-862B-65

Query Match 0.4%; Score 78.4; DB 4; Length 105;
Best Local Similarity 84.6%; Pred. No. 6.9e-09;
Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Oy 15129 TTTTGTATTTTGTAGAGATGGAGTTTCGCCGTTAGCCAGGATGCTCTCGATCTCC 15188
Db 2 TTTTGTATTTTGTAGAGATGGAGTTTCGCCGTTAGCCAGGATGCTCTCGATCTCC 61

Oy 15189 TCACCTCTGTATCCACCGCCCTCGCCCTCCCAAAAGTCTGGGAT 15232
Db 62 TGACCTTGTATCCACCGCCCTCGCCCTCCCAAAAGTCTGGGAT 105

RESULT 5
US-08-787-739-65
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/787,739
FILING DATE: 24-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/486,756
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/481,658
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,862
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,863
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,077
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.4
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-787-739-65

Query Match 0.4%; Score 78.4; DB 5; Length 105;
Best Local Similarity 84.6%; Pred. No. 6.9e-09;
Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Oy 15129 TTTTGTATTTTGTAGAGATGGAGTTTCGCCGTTAGCCAGGATGCTCTCGATCTCC 15188
Db 2 TTTTGTATTTTGTAGAGATGGAGTTTCGCCGTTAGCCAGGATGCTCTCGATCTCC 61

Oy 15189 TCACCTCTGTATCCACCGCCCTCGCCCTCCCAAAAGTCTGGGAT 15232
Db 62 TGACCTTGTATCCACCGCCCTCGCCCTCCCAAAAGTCTGGGAT 105

RESULT 6
US-08-481-658B-65/c
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:

MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-486-756A-65

Query Match 0.4%; Score 72.8; DB 4; Length 105;
Best Local Similarity 86.8%; Pred. No. 1.2e-07;
Matches 92; Conservative 0; Mismatches 12; Indels 2; Gaps 1;

QY 10402 ATCCGAGCACTTTGGCAGGCTGAGCGGCGCAGATCATTGAGTCAGGAGTTTGAGACCA 10461
|||||
DB 105 ATCCGAGCACTTTGGGAGCGCCGAGCTGGTGATCAC--AAGTCAGGAGTTTGAGAGCA 48
|||||
QY 10462 GCCTGCCAACATGTGTAACCCCTGTCTCCACTAAAAATACAAAA 10507
|||||
DB 47 GCCTGCCAATATGTTGAACCCCTGTCTCTACTAAGATGTAAAA 2
|||||

RESULT 9
US-08-485-862B-65/c
; Sequence 65, Application US/08485862B
; Patent No. 5989838
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,862B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727

INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-485-862B-65

Query Match 0.4%; Score 72.8; DB 4; Length 105;
Best Local Similarity 86.8%; Pred. No. 1.2e-07;
Matches 92; Conservative 0; Mismatches 12; Indels 2; Gaps 1;

QY 10402 ATCCGAGCACTTTGGCAGGCTGAGCGGCGCAGATCATTGAGTCAGGAGTTTGAGACCA 10461
|||||
DB 105 ATCCGAGCACTTTGGGAGCGCCGAGCTGGTGATCAC--AAGTCAGGAGTTTGAGAGCA 48
|||||
QY 10462 GCCTGCCAACATGTGTAACCCCTGTCTCCACTAAAAATACAAAA 10507
|||||
DB 47 GCCTGCCAATATGTTGAACCCCTGTCTCTACTAAGATGTAAAA 2
|||||

RESULT 10
US-08-787-739-65/c
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332

INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-787-739-65


```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340.426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; US-08-340-426D-91

Query Match 0.4%; Score 63.8; DB 4; Length 84;
Best Local Similarity 85.5%; Pred. No. 1.1e-05;
Matches 71; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 15160 CCGTGTACCGAGGATGGTCTCGATCTCTGACCTCGTGATCCACGGGCTCGGCCTCCC 15219
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1 CCATGTTTCATCAGGCTGGTGTGCAACTCTGACCTCGTGATCCGCGCTCAGCCTCCC 60

QY 15220 AAAGTCTGGGATTACAGCGTG 15242
||||||| ||||| ||||| |||||
Db 61 AAAGTCTGGGATTACAGCGTG 83

RESULT 15
PCT-US95-17111A-91
; Sequence 91, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/17111A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,426
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; PCT-US95-17111A-91

Query Match 0.4%; Score 63.8; DB 6; Length 84;
Best Local Similarity 85.5%; Pred. No. 1.1e-05;
Matches 71; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 15160 CCGTGTACCGAGGATGGTCTCGATCTCTGACCTCGTGATCCACGGGCTCGGCCTCCC 15219
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1 CCATGTTTCATCAGGCTGGTGTGCAACTCTGACCTCGTGATCCGCGCTCAGCCTCCC 60

QY 15220 AAAGTCTGGGATTACAGCGTG 15242
||||||| ||||| ||||| |||||
Db 61 AAAGTCTGGGATTACAGCGTG 83

RESULT 14
US-08-450-673C-91
; Sequence 91, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; US-08-450-673C-91

Query Match 0.4%; Score 63.8; DB 4; Length 84;
Best Local Similarity 85.5%; Pred. No. 1.1e-05;
Matches 71; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 15160 CCGTGTACCGAGGATGGTCTCGATCTCTGACCTCGTGATCCACGGGCTCGGCCTCCC 15219
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1 CCATGTTTCATCAGGCTGGTGTGCAACTCTGACCTCGTGATCCGCGCTCAGCCTCCC 60

QY 15220 AAAGTCTGGGATTACAGCGTG 15242
||||||| ||||| ||||| |||||
Db 61 AAAGTCTGGGATTACAGCGTG 83

RESULT 15
PCT-US95-17111A-91
; Sequence 91, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/17111A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,426
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; PCT-US95-17111A-91

Query Match 0.4%; Score 63.8; DB 6; Length 84;
Best Local Similarity 85.5%; Pred. No. 1.1e-05;
Matches 71; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 15160 CCGTGTACCGAGGATGGTCTCGATCTCTGACCTCGTGATCCACGGGCTCGGCCTCCC 15219
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1 CCATGTTTCATCAGGCTGGTGTGCAACTCTGACCTCGTGATCCGCGCTCAGCCTCCC 60

QY 15220 AAAGTCTGGGATTACAGCGTG 15242
||||||| ||||| ||||| |||||
```

Db 61 AACTGTGCTGGATTACAAGCGTG 83

Search completed: June 20, 2000, 18:07:26
Job time: 537176 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2000, 17:57:22 ; Search time 15086.9 Seconds
(without alignments)
-1568.587 Million cell updates/sec

Title: US-08-852-495C-2_COPY_213000_237326

Perfect score: 24327

Sequence: 1 TAGAGTTAAATGTGAAAAAT.....TGTGTGTGTGTGTGTGTGTG 24327

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_bal.*

2: gb_ba2.*

3: gb_om.*

4: gb_ov.*

5: gb_pat.*

6: gb_ph.*

7: gb_pl1.*

8: gb_pl2.*

9: gb_pr1.*

10: gb_pr2.*

11: gb_pr3.*

12: gb_ro.*

13: gb_sts.*

14: gb_sy.*

15: gb_un.*

16: gb_v1.*

17: em_fun.*

18: em_hum1.*

19: em_hum2.*

20: em_in.*

21: em_om.*

22: em_or.*

23: em_ov.*

24: em_pat.*

25: em_ph.*

26: em_pl.*

27: em_ro.*

28: em_sts.*

29: em_sy.*

30: em_un.*

31: em_v1.*

32: gb_htg1.*

33: gb_htg2.*

34: gb_in1.*

35: gb_in2.*

36: em_ba1.*

37: em_ba2.*

38: em_hum3.*

39: em_hum4.*

40: gb_pr4.*

41: gb_htg3.*

42: gb_htg4.*

43: gb_htg5.*

44: gb_htg6.*

45: gb_htg7.*
46: em_htg1.*
47: em_htg2.*
48: em_htg3.*
49: em_hum5.*
50: gb_pl3.*
51: gb_pr5.*
52: gb_htg8.*
53: gb_htg9.*
54: gb_htg10.*
55: gb_htg11.*
56: gb_htg12.*
57: gb_htg13.*
58: gb_htg14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	98.4	0.4	108	10	HSIDLNRN2	X05250 Human LDL-r
C 2	90.8	0.4	107	9	HUMALCE162	M87924 Human carc
C 3	90.4	0.4	108	10	HSIDLNRN2	X05250 Human LDL-r
C 4	89.6	0.4	108	11	HSU67803	U67803 Human small
C 5	87.8	0.4	108	10	HSIDLNRD1	X05249 Human LDL-r
C 6	87.8	0.4	108	10	HSIDLNRD2	X05251 Human LDL-r
C 7	87.2	0.4	103	9	HUMALCE221	M87896 Human carc
C 8	82.4	0.3	104	9	HUMALCE272	M87899 Human carc
C 9	81.4	0.3	108	10	HSIDLNRD1	X05249 Human LDL-r
C 10	81.4	0.3	108	10	HSIDLNRD2	X05251 Human LDL-r
C 11	81.6	0.3	108	11	HSU67804	U67804 Human small
C 12	79.6	0.3	107	9	HUMALCE162	M87924 Human carc
C 13	78.8	0.3	103	9	HUMALCE221	M87896 Human carc
C 14	78.8	0.3	110	9	HUMALCE43	M87900 Human carc
C 15	76.8	0.3	108	10	HSIDLNRD1	X05248 Human LDL-r
C 16	76.2	0.3	101	10	S79560	S79560 HRX (Intron
C 17	75.8	0.3	108	9	HUMD1D03M5	D16965 Human HepG2
C 18	75	0.3	103	13	HS8IC8R	X57789 Human sequ
C 19	75.2	0.3	108	11	HSU67808	U67808 Human small
C 20	74.4	0.3	110	11	HSU67807	U67807 Human small
C 21	73.6	0.3	106	13	G32743	G32743 A009P31 Hum
C 22	72.4	0.3	91	13	HUMUT8164A	L30244 Human STS U
C 23	72.6	0.3	108	9	HUMD1D03M5	D16965 Human HepG2
C 24	72	0.3	90	9	HUMDLRFL	K03555 Human low d
C 25	70.4	0.3	107	11	HSU67806	U67806 Human small
C 26	70.4	0.3	108	10	HSIDLNRD1	X05248 Human LDL-r
C 27	69.8	0.3	102	13	G32906	G32906 A009W09 Hum
C 28	69.4	0.3	100	13	HUMUT931A	L31299 Human STS U
C 29	69.6	0.3	104	9	HUMALCE272	M87899 Human carc
C 30	68.4	0.3	79	10	S73203	S73203 ALI-1 (tand
C 31	68.4	0.3	95	13	HUMUT8002B	L30176 Human STS U
C 32	67	0.3	84	5	AR051521	AR051521 Sequence
C 33	66.8	0.3	99	13	HUMUT7692A	L30306 Human STS U
C 34	67	0.3	108	13	G43535	G43535 WIAF-2393-S
C 35	66	0.3	91	13	HUMUT8164A	L30244 Human STS U
C 36	66.2	0.3	100	9	HUMGALNS	D45223 Human GALNS
C 37	65.6	0.3	80	9	HUMBRKFAE	M36135 Human alpha
C 38	65.8	0.3	97	9	HUMDLR2A	M14180 Human low d
C 39	65.6	0.3	107	13	G32919	G32919 A009W27 Hum
C 40	65.2	0.3	94	9	HUMHGAL	M13479 Human alpha
C 41	65.2	0.3	95	10	HSSTHPKIB	X66361 H.sapiens m
C 42	65	0.3	98	13	G33095	G33095 EYRP13C9R H
C 43	65	0.3	100	13	HUMUT931A	L31299 Human STS U
C 44	64	0.3	80	9	HUMBRKFAE	M36135 Human alpha
C 45	63.8	0.3	106	13	G32743	G32743 A009P31 Hum

ALIGNMENTS

```

RESULT 1
HSLDLRN2/c
LOCUS HSLDLRN2 108 bp DNA PRI 20-MAY-1992
DEFINITION Human LDL-receptor gene Intron 14 fragment (normal gene).
ACCESSION X05250
VERSION X05250.1 GI:34337
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
Williamson,R. and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901
COMMENT See X05252 for deletion junction
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES
source
Location/Qualifiers
1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
intron
1..108
/note="intron XIV fragment"
BASE COUNT 28 a 23 c 39 g 18 t
ORIGIN
Query Match 0.4%; Score 98.4; DB 10; Length 108;
Best Local Similarity 94.4%; Pred. No. 1.9e-05;
Matches 102; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 16893 CTGGCTCAGTGCACCTCTGCTCTGGGTCAAGCAATTCCTGCTCAGCTTCCA 16942
|||
Db 108 CTGGCTCAGTGCACCTCTGCTCTGGGTCAAGCAATTCCTGCTCAGCTCCTCCG 49
|||

Qy 16943 AGTAGCTGGATTACAGCAGCCACTACACGCTGGCTAATTTTGT 16990
|||||
Db 48 AGTAGCTGGATTACAGCAGCTGCGCACCAGCTGGCTAATTTTGT 1

RESULT 2
HUMALCE162/c
LOCUS HUMALCE162 107 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE162.
ACCESSION M87924
VERSION M87924.1 GI:174871
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 107)
AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
source
Location/Qualifiers
1..107
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="NTERa2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT 28 a 30 c 35 g 14 t
ORIGIN

```

```

Query Match 0.4%; Score 90.8; DB 9; Length 107;
Best Local Similarity 93.1%; Pred. No. 0.00025;
Matches 95; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 14072 TTTTGTGACGAGGTCCTCCTGTGTACCCAGGCTGGAGTGCAGTGGCGGCGATTCGGCT 14131
|||||
Db 106 TTTTGTGACGAGGTCCTCCTGTGTGTGCCAGGCTGGAGTGCAGTGGCGGCGATTCGGCT 47
|||||

Qy 14132 CACTGCAACCTCCGCTCCCGGTTCAAGTCATCTCCTGCC 14173
|||||
Db 46 CACTGCAACCTCCGCTCCCGGTTCAAGTCATCTCCTGCC 5

RESULT 3
HSLDLRN2
LOCUS HSLDLRN2 108 bp DNA PRI 20-MAY-1992
DEFINITION Human LDL-receptor gene Intron 14 fragment (normal gene).
ACCESSION X05250
VERSION X05250.1 GI:34337
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
Williamson,R. and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901
COMMENT See X05252 for deletion junction
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES
source
Location/Qualifiers
1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
intron
1..108
/note="intron XIV fragment"
BASE COUNT 28 a 23 c 39 g 18 t
ORIGIN
Query Match 0.4%; Score 90.4; DB 10; Length 108;
Best Local Similarity 89.8%; Pred. No. 0.00028;
Matches 97; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 7243 ACAAAATAGCCAGGTGTGTGGCAGGCACCTGTATCCAGCTATTCAGAGGCTGAG 7302
|||||
Db 1 ACAAAATAGCCAGGTGTGTGGCAGGTGCTGTATCCAGCTACTCTCGGAGGCTGAG 60
|||||

Qy 7303 ACAGGAGATCGCTTGACCCAGGAGGTGGAGTTGCATTGAGCCAAAG 7350
|||||
Db 61 GCAGGAGAAATTCCTTGAAACCCAGGAGGAGGTTCAGTGAGCGCGAG 108
|||||

RESULT 4
HSU67803/c
LOCUS HSU67803 108 bp RNA PRI 01-AUG-1997
DEFINITION Human small cytoplasmic Alu transcript.
ACCESSION U67803
VERSION U67803.1 GI:2289917
KEYWORDS Alu.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE cDNAs derived from primary and small cytoplasmic Alu (scAlu)
transcripts
JOURNAL J. Mol. Biol. 271 (2), 222-234 (1997)

```


Db 6 CCTGTAATCCAGCTACACGGAAGCTAAGCAGGAGTAATCGTTGAACCGGGAGCGG 65
Qy 7333 AGGTTGCATTGAGCAAGATCATGCACTGCACCTCCAG 7370
Db 66 AGTTGCAGTGAAGCGGAGATCGTGCCATTGCACCTCCAG 103

RESULT 14
HUMALCE43/c HUMALCE43 110 bp ss-RNA PRI 15-APR-1994
LOCUS Human carcinoma cell-derived Alu RNA transcript, clone CE43.
DEFINITION M87900
ACCESSION M87900.1 GI:174876
VERSION
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 110)
AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) in press
FEATURES
Location/Qualifiers
source 1..110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Wtera2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT 27 a 31 c 34 g 18 t
ORIGIN

Query Match 0.3%; Score 78.8; DB 9; Length 110;
Best Local Similarity 84.0%; Pred. No. 0.014;
Matches 89; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
Qy 12741 GTTTCACCATGTAGCCAGGATGGTCTCGATCTCTGACCTCGTGATCCGCCACCTGAG 12800
Db 110 GTTTCGTCATGTAGCCAGGATGGTCTTGAACACTAGCTGCAATCTCTGCGCTGG 51
Qy 12801 CCTCCAAAGTGGGATACAGCTGTGAGCCAGCGCGCGGCC 12846
Db 50 CCTCCAAAGTGGGATGTAGTGTGAGCCAGCGCGCGGCC 5

RESULT 15
HSLDL12 HSLDL12 108 bp DNA PRI 20-MAY-1992
LOCUS Human LDL-receptor gene intron 12 fragment (normal gene) LBL - low
DEFINITION density lipoprotein.
ACCESSION X05248
VERSION X05248.1 GI:34334
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor;
repetitive sequence.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
Williamson,R. and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901
COMMENT see X05249 for deletion junction
DATA kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES Location/Qualifiers

source 1..108
misc_feature /db_xref="taxon:9606"
complement(<1..65)
/note="Alu repeat"
intron 1..108
/note="intron XII fragment"
BASE COUNT 21 a 38 c 20 g 29 t
ORIGIN
Query Match 0.3%; Score 76.8; DB 10; Length 108;
Best Local Similarity 83.7%; Pred. No. 0.027;
Matches 87; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
Qy 20565 TCGGCTCACAGCAACCTCCACCTCTCTGGGTTCAAGTGATTTCTCTGCCTCAGCCTCCTGA 20624
Db 2 TCGGCTCACAGCAACCTCTCTGGGTTCAAGTGATTTCTCTGCCTCAGCCTCCTGA 61
Qy 20625 GTAGCTGGGATTACAGAGCTGTGTCAACACACACCTGGCTAAATTTT 20668
Db 62 GTAGCTGGGATTACAGAGCTGTGTCAACACACACCTGGCTAAATTTT 105
Search completed: June 21, 2000, 05:17:36
Job time: 578340 sec


```
RESULT 2
ID T24892/c
AC T24892;
DE T24892 standard; cDNA to mRNA; 100 BP.
DE Human gene signature HUMGS06998.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1720; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;

Query Match 0.3%; Score 67.2; DB 1; Length 100;
Best Local Similarity 78.8%; Pred. No. 0.24;
Matches 78; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 14066 TATTATTATTTTCAGACGAGTCTCACTGTGCACCCAGCGCTGAGTGCGCGGATC 14125
DB 100 TTTGTGTTGTTTCAACAGAGTGTCACTGTGTCACCCAGCGGAGTGCAANGTGCAATC 41

QY 14126 TCGGCTCACTGCAACCTCCGCTCCCGGGTTCAAGTGAT 14164
DB 40 TCAGCTNATTGCAAAATCTGCTCCCGAGTTCACGCGAT 2

RESULT 3
ID T26828
AC T26828 standard; cDNA to mRNA; 108 BP.
DE T26828 (first entry)
DE Human gene signature HUMGS09078.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
```

```
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 2182; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 18 A; 33 C; 23 G; 28 T;

Query Match 0.3%; Score 64.4; DB 1; Length 108;
Best Local Similarity 89.5%; Pred. No. 0.56;
Matches 68; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3995 ATCTCCTGACCTCGGATCCGCCGCTCCGCTCCCAAGTGGGATTTACAGCGTG 4054
DB 2 ATCTCCTGACCTCGGATCCGCCGCTCCGCTCCCAAGTGGGATTTACAGCGATG 61

QY 4055 AGCCACCGTGCCTGCGC 4070
DB 62 AGCCACCGTGCCTGCGC 77

RESULT 4
ID T25009/c
AC T25009 standard; cDNA to mRNA; 108 BP.
DE T25009 (first entry)
DE Human gene signature HUMGS07131.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1748; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
```

```

SQ Sequence 108 BP; 34 A; 31 C; 26 G; 15 T;

Query Match
Best Local Similarity 0.3%; Score 64; DB 1; Length 108;
Matches 79; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 285 TTTGTTGTTGTTTCTTTTTCAGATAGAGTCTCTCTGTCATTCAGCGCTGGAGTGCAGTGG 344
   || | ||||| ||| | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 107 TTGTTGTTGTTGTTTTCACAGGGTCTGCTGTCATTCAGCGCTGGAATNCAGTGG 48

QY 345 CATGATCTCAGCTCAGTGCAGCTCCGCTCCCGGGTTCAAGAGAT 390
   || | | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 47 CGTGACCATGCTCATGTCAGCGCTTGCCCTCATGGGCTCAGGCGAT 2

RESULT 5
ID T25009 standard; cDNA to mRNA; 108 BP.
AC T25009;
DE Human gene signature HUMGS07131.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
PI WPI: 95-206931/27.
DR Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1748; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 34 A; 31 C; 26 G; 15 T;

Query Match
Best Local Similarity 0.3%; Score 63.4; DB 1; Length 108;
Matches 79; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 7311 ATCGCTTGAACCAAGGAGTGGAGTTGATTCAGCAAGATCATGCCACTGCACATCCAG 7370
||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 ATCGCTGAGCCCAAGGAGGCAAGGCTGCAGTGCATGAGCCATGGTCACGCCACTGNATCCAG 61

QY 7371 CTTGGCGGCACAGAGTAAGACTCCGTTTCAAAAACAAAAACCCCAAGA 7417
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 CTTGAGTGACAGCAAGACCCCTGTTTGAACAACAACAACAAACAA 108

RESULT 6
T25854/c
ID T25854 standard; cDNA to mRNA; 91 BP.

Query Match
Best Local Similarity 0.3%; Score 62.6; DB 1; Length 91;
Matches 71; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 14076 TGAGACGAGTCTCAGTCTCTCACCAGGCTGGAGTGCAGTGGCGGTCACT 14135
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 90 TGAGACAGNNCTCAGCGCTGTACACNAGGCTGGAGGCGCAGGAGTCCATCTCAGCTCACT 31

QY 14136 GCAACCTCCGCTCCCGGGTTCAAGTGAT 14164
||||| ||||| ||||| ||||| |||||
Db 30 TGAACCCNCTGCCCTCCCTAGGCTCAAGTGAT 2

RESULT 7
T24892
ID T24892 standard; cDNA to mRNA; 100 BP.
AC T24892;
DE Human gene signature HUMGS06998.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
PI WPI: 95-206931/27.
DR Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
```

PS Claim 1; Page 1720; 2245pp; Japanese.

CC A single-stranded DNA (or its complementary strand or the corresp.

CC double-stranded DNA) which comprises one of the 7837 "GS" sequences

CC given in T19001-T26837 and which is able to hybridise to part of

CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)

CC sequences were obtained from 3'-directed cDNA libraries prepared

CC from various human tissues; synthesis of cDNA was initiated from the

CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-

CC untranslated sequence is unique to a particular mRNA species, almost

CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library

CC is constructed so as to reflect accurately the relative abundance of

CC different mRNAs in the particular tissue from which it was derived.

CC The appearance frequency of a given GS in a cDNA library can be

CC determined (esp. using primers and probes derived from the GS

CC sequences) as a means of diagnosing abnormal cell function or for

CC recognising different cell types.

CC Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;

SQ

Query Match 0.2%; Score 60.8; DB 1; Length 100;

Best Local Similarity 74.7%; Pred. No. 1.7;

Matches 74; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 7311 ATCGTGTGAACCCAGGAGTGGAGTTGCATTTGAGCAAGATCATGCACCTGCACCTCCAG 7370

Db 2 ATCGTGTGAACCTGGGAGGAGCAATTTGCAATNAGCTGAGATTGCACCTTGCACTCCNG 61

QY 7371 CTTGGGGCAGCAGTAGACATCTCGTTTCAAAAACAAAA 7409

Db 62 CTTGGGTGCAGAGTAGTCACTCTTTTGAACAACAAACA 100

RESULT 8

T26213

ID T26213 standard; cDNA to mRNA; 103 BP.

AC T26213;

DE 13-NOV-1996 (first entry)

DT Human gene signature HUMGS08452.

KE Gene signature; messenger RNA; mRNA; relative abundance; frequency;

KW human; cloning; mapping; non-biased library; diagnosis; detection;

KW cell typing; abnormal cell function; ss.

OS Homo sapiens.

PN WO9514772-A1.

PD 01-JUN-1995.

PF 11-NOV-1994; J01916.

PR 12-NOV-1993; JP-355504.

PA (MATS/) MATSUBARA K.

PA (OKUB/) OKUBO K.

PI Matsubara K, Okubo K;

PI WPI; 95-206931/27.

DR Identifying gene signatures in 3'-directed human cDNA library - e.g.

PT for diagnosis of abnormal cell function, by preparing cDNA that

PT reflects relative abundance of corresp. mRNA in specific human

PT tissues

PS Claim 1; Page 2029; 2245pp; Japanese.

CC A single-stranded DNA (or its complementary strand or the corresp.

CC double-stranded DNA) which comprises one of the 7837 "GS" sequences

CC given in T19001-T26837 and which is able to hybridise to part of

CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)

CC sequences were obtained from 3'-directed cDNA libraries prepared

CC from various human tissues; synthesis of cDNA was initiated from the

CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-

CC untranslated sequence is unique to a particular mRNA species, almost

CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library

CC is constructed so as to reflect accurately the relative abundance of

CC different mRNAs in the particular tissue from which it was derived.

CC The appearance frequency of a given GS in a cDNA library can be

CC determined (esp. using primers and probes derived from the GS

CC sequences) as a means of diagnosing abnormal cell function or for

CC recognising different cell types.

CC Sequence 103 BP; 33 A; 21 C; 25 G; 23 T;

SQ

KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-AL.

PD 01-JUN-1995.

PF 11-NOV-1994; J01916.

PR 12-NOV-1993; JP-355504.

PA (MATS/) MATSUBARA K.

PA (OKUB/) OKUBO K.

PI Matsubara K, Okubo K;

PI WPI; 95-206931/27.

DR Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 2039; 2245pp; Japanese.

CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.

CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

SQ Sequence 103 BP; 33 A; 21 C; 25 G; 23 T;

Query Match 0.2%; Score 58.4; DB 1; Length 103;

Best Local Similarity 74.0%; Pred. No. 3.5;

Matches 74; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 14065 TTATTTATTTTTCAGCGGAGTCTACTCTGCACCGCTGCAGTGGCGCGAT 14124

DB 101 TTTTTCCTTAAAGACATGTTCTTACTCTGTGGCGGCTGCAGTGGCGCCAT 42

QY 14125 CTCGGCTCACTGCAACCTCGCCCTCCGGGTTCAAGTGAT 14164

DB 41 CATAGCTCACTGTAAACACCAACTCTGCACATCAAGTGAT 2

RESULT 11

ID T20927 standard; cDNA to mRNA; 103 BP.

AC T20927;

DT 24-JUL-1996 (first entry)

DE Human gene signature HUMGS0180.

KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;

KW human; cloning; mapping; non-biased library; diagnosis; detection;

KW cell typing; abnormal cell function; ss.

OS Homo sapiens.

PN WO9514772-AL.

PD 01-JUN-1995.

PF 11-NOV-1994; J01916.

PR 12-NOV-1993; JP-355504.

PA (MATS/) MATSUBARA K.

PA (OKUB/) OKUBO K.

PI Matsubara K, Okubo K;

PI WPI; 95-206931/27.

DR Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 758-759; 2245pp; Japanese.

CC A single-stranded DNA (or its complementary strand or the corresp.

CC double-stranded DNA) which comprises one of the 7837 "GS" sequences

CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.

CC The appearance frequency of a given GS in a cDNA library can be

CC determined (esp. using primers and probes derived from the GS

CC sequences) as a means of diagnosing abnormal cell function or for

CC recognising different cell types.

SQ Sequence 103 BP; 22 A; 27 C; 31 T;

Query Match 0.2%; Score 58; DB 1; Length 103;

Best Local Similarity 73.0%; Pred. No. 4;

Matches 73; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 20601 GATTCCTCCCTCAGCCCTCCTCAGTAGCTGGGATTACAGACCTGTGTACACACCTGG 20660

DB 1 GATCCCTCCCACTTCCACCTCCCAAGTAGCTGTGGCTGTGTGCCACCATGTCCAG 60

QY 20661 CTAATTTTGTATTTTAGTAGACAGACAGGGTTTCACCGTG 20700

DB 61 CTGATTTTNGTATTTTNTAGTAGGACAGATATTTCTCCATG 100

RESULT 12

T20373/c

ID T20373 standard; cDNA to mRNA; 107 BP.

AC T20373;

DT 19-JUL-1996 (first entry)

DE Human gene signature HUMGS01525.

KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;

KW human; cloning; mapping; non-biased library; diagnosis; detection;

KW cell typing; abnormal cell function; ss.

OS Homo sapiens.

PN WO9514772-AL.

PD 01-JUN-1995.

PF 11-NOV-1994; J01916.

PR 12-NOV-1993; JP-355504.

PA (MATS/) MATSUBARA K.

PA (OKUB/) OKUBO K.

PI Matsubara K, Okubo K;

PI WPI; 95-206931/27.

DR Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 623; 2245pp; Japanese.

CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the

CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.

CC The appearance frequency of a given GS in a cDNA library can be

CC determined (esp. using primers and probes derived from the GS

CC sequences) as a means of diagnosing abnormal cell function or for

CC recognising different cell types.

SQ Sequence 107 BP; 26 A; 29 C; 17 G; 29 T;

Query Match

Best Local Similarity 72.6%; Pred. No. 6.5;

Matches 69; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2000, 09:51:36 ; Search time 7160.88 Seconds
(without alignments)
13769.643 Million cell updates/sec

Title: US-08-852-495C-2_COPY_213000_237326
Perfect score: 24327
Sequence: 1 TAGAGTTAAATGTGAAAAAT.....TGTGTGTGTGTGTGTGTG 24327

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : EST:
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	96.4	0.4	106	37	AA703692	ag81a10.r
2	94.8	0.4	109	30	AA243009	zr25h02.s
3	92.4	0.4	103	38	AA807640	nx08b05.s
4	92.6	0.4	108	84	B65160	CIT-HSP-201
5	92.4	0.4	110	39	AA897366	am06h02.s
6	91.4	0.4	109	94	AQ028426	CIT-HSP-2
7	91	0.4	108	84	B65160	CIT-HSP-201
8	90.4	0.4	110	30	AA244245	nc07a04.s
9	89.8	0.4	101	39	AA855205	ak64h01.s
10	88.4	0.4	106	105	AQ264176	CITBI-EI-
11	87.8	0.4	103	84	B48914	RPCI11-4A12
12	87.8	0.4	107	35	AA565533	nk42b11.s
13	87	0.4	103	30	AA228795	nc14e07.s
14	86.8	0.4	103	108	AQ582186	RPCI11-4
15	87	0.4	103	108	AQ584425	RPCI11-4
16	86.6	0.4	103	108	AQ535244	RPCI11-3
17	86.8	0.4	106	38	AA812141	ob48h02.s
18	86.4	0.4	105	30	AA218889	zql5d04.s
19	86.4	0.4	107	39	AA828124	od71a07.s
20	86.6	0.4	109	94	AQ028426	CIT-HSP-2
21	86.6	0.4	110	30	AA244245	nc07a04.s
22	85.6	0.4	102	30	AA228656	nc19f09.s
23	85.8	0.4	105	28	AA078003	7H12D08 C
24	85.8	0.4	105	105	AQ282107	RPCI11-94
25	85.2	0.4	106	94	AQ062963	CIT-HSP-2
26	85.2	0.4	109	22	H1143	YM08c06.r1
27	84.8	0.3	104	29	AA129957	zn86h04.r
28	84.8	0.3	104	29	AA129957	zn86h04.r
29	84.8	0.3	107	33	AA385808	EST99495
30	85	0.3	109	84	B17434	345K2.TVB C
31	85	0.3	109	105	AQ265749	CITBI-EI-
32	84.6	0.3	107	103	AQ240182	CIT-HSP-2
33	83.8	0.3	108	94	AQ014433	CIT-HSP-2
34	83.4	0.3	101	39	AA835205	ak64h01.s
35	83.4	0.3	102	94	AQ004934	CIT-HSP-2
36	83.6	0.3	109	24	N25299	YW52c09.s1
37	83	0.3	101	33	AA381369	EST94442
38	83.2	0.3	104	108	AQ544583	CITBI-EI-
39	83.2	0.3	106	106	AQ418993	RPCI11-1
40	83.2	0.3	107	24	N23686	YW46a02.s1
41	83.4	0.3	110	94	AQ003188	RPCI11-1D
42	82.6	0.3	106	108	AQ544957	CITBI-EI-
43	82.8	0.3	110	29	AA177157	nc02g07.s
44	82	0.3	98	24	H67549	yu68f10.s1
45	82.2	0.3	104	105	AQ268072	RPCI11-73

ALIGNMENTS

```

RESULT 1
AA703692 106 bp mRNA EST 24-DEC-1997
LOCUS ag81a10.r1 Stratagene hnt neuron (#937233) Homo sapiens cDNA clone
DEFINITION IMAGE:1140858 5' similar to contains Alu repetitive element; mRNA
sequence.
ACCESSION AA703692
VERSION AA703692.1 GI:2713610
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 106)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geissel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1397630.

```

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 53.

FEATURES

Location/Qualifiers
1..106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1140858"
/clone_lib="Stratagene hnt neuron (#937233)"
/dev_stage="hnt neurons"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT.
Differentially, post mitotic hnt neurons. Average insert
size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'
GAATTCGGCAGGAG 3' -3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTT 3'

BASE COUNT

19 a 29 c 29 g 29 t

Query Match

Best Local Similarity 0.4%; Score 96.4; DB 37; Length 106;
Matches 100; Conservative 94.3%; Pred. No. 0.14;
Mismatches 6; Indels 0; Gaps 0;

QY 12724 TTTTGTAGACAGCGGGTTTCCACCATGTTAGCCAGGATGCTCGATCTCTGACCTCG 12783

Db 1 TTTTGTAGACAGCGGGTTTCCACCATGTTAGCCAGGATGCTCGATCTCTGACCTCG 60

QY 12784 TGATCGCCACCTGAGCTCCCAAGTCTGGGATTACAGGTGTG 12829

Db 61 TGATCGCCGCTCAGCTCCCAAGTCTGGGATTACAGGCGTG 106

RESULT 2

AA243009 109 bp mRNA EST 11-MAR-1998
LOCUS zr25h02.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
DEFINITION cDNA clone IMAGE:664467 3' similar to contains Alu repetitive
element; contains element LTR1 repetitive element; mRNA sequence.

AA243009 109 bp mRNA EST 11-MAR-1998

VERSION AA243009.1 GI:1873869

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 109)

AUTHORS

Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geissel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.

WashU-NCI human EST Project

TITLE

Unpublished (1997)

JOURNAL

On Dec 3, 1996 this sequence version replaced gi:1126869.

COMMENT

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1127 Std Error: 0.00

Seq primer: -41ml3 fwd. ET from Amersham

High quality sequence stop: 102.

FEATURES
source

Location/Qualifiers
1. .103
/organism="Homo sapiens"
/db_xref="GDB:5426481"
/db_xref="taxon:9606"
/clone="IMAGE:664467"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/tissue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial cells"
/lab_host="SOUR (kanamycin resistant)"
/note="Organ: brain; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
oligo dt. Uninduced, exponentially growing neuroepithelial
cells (Ntera-2/cl.D1). Average insert size: 1.0 kb;
Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCAGGAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3"

19 a 30 c 30 g 30 t

BASE COUNT
ORIGIN

Query Match 0.4%; Score 94.8; DB 30; Length 109;
Best Local Similarity 93.4%; Pred. No. 0.2;
Matches 99; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 12724 TTTTGTAGACAGGGGTTTCACCATGTTAGCCAGGATGCTCGATCTCCTGACCTCG 12783
|||||
Db 4 TTTTGTAGACAGGGGTTTCACCGTGTAGCCAGGATGCTTTCATCTCCTACCTCG 63
|||||

QY 12784 TGATCGGCCACCTGAGCCTCCCAAGTCTCGGATTCAGAGTG 12829
|||||
Db 64 TGATCGGCCACCTGCGCTCCCAAGTCTCGGATTCAGAGCGTG 109
|||||

RESULT 3
AA807640

LOCUS AA807640 103 bp mRNA EST 05-MAR-1998
DEFINITION nx08b05.s1 NCI_CGAP_GC3 Homo sapiens cDNA clone IMAGE:1255473 3'
similar to contains Alu repetitive element;; mRNA sequence.

ACCESSION AA807640
VERSION AA807640.1 GI:2877108
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 103)
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
TITLE Unpublished (1997)
JOURNAL On Jan 19, 1998 this sequence version replaced gi:2151346.
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
Cloning Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 774 Std Error: 0.00
Seq primer: -40ml3 fwd. Et from Amersham
High quality sequence stop: 87.

FEATURES
source

Location/Qualifiers
1. .103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1255473"
/clone_lib="NCI_CGAP_GC3"
/tissue_type="pooled germ cell tumors"

/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is not normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo."

19 a 27 c 30 g 27 t

BASE COUNT
ORIGIN

Query Match 0.4%; Score 92.4; DB 38; Length 103;
Best Local Similarity 94.1%; Pred. No. 0.38;
Matches 96; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 12729 AGTAGAGCGGGGTTTCACCATGTTAGCCAGGATGCTCGATCTCCTGACCTCGTGATC 12788
|||||
Db 2 AGTAGAGATGGGGTTTCACCGTGTAGCCAGGATGCTCGATCTCCTGACCTTGTGATC 61
|||||

QY 12789 CGCCACCTGAGCCTCCCAAGTCTGGGATTACAGGTGTA 12830
|||||
Db 62 CGCTCACCTCGGCTCCCAAGTCTGGGATTACAGGTGTA 103
|||||

RESULT 4

B65160 108 bp DNA GSS 21-JUN-1998
LOCUS CIT-HSP-2017G2.TRB CIT-HSP Homo sapiens genomic clone 2017G2,
genomic survey sequence.

ACCESSION B65160
VERSION B65160.1 GI:2639138
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 108)
REFERENCE Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
AUTHORS Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: CIT-HSP-2017G2.TFB
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1. .108
/organism="Homo sapiens"
/db_xref="GDB:7043860"
/db_xref="taxon:9606"
/clone="2017G2"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
HindIII"

26 a 27 c 34 g 21 t

BASE COUNT
ORIGIN

REFERENCE 1 (bases 1 to 108)
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other GSSs: CIT-HSP-2017G2.TFB
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source
1..108
/organism="Homo sapiens"
/db_xref="GDB:7043860"
/db_xref="taxon:9606"
/clone="2017G2"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"

BASE COUNT 26 a 27 c 34 g 21 t
ORIGIN
Query Match 0.4%; Score 91; DB 84; Length 108;
Best Local Similarity 90.7%; Pred. No. 0.52;
Matches 97; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 20547 GAGTCAGTGGGATGATCTCGGCTACACACACCTCCACCTCCTGGTTCAAGTGATTTCT 20606
|||||
Db 107 GTGTGAGTGGTATGATCTGTGGCTACTGCAACCTCCACCTCCCGGGTTCAAGAGATTTCT 48
|||||

QY 20607 CCTGCTCAGCTCTCTGAGTAGTGGATTACAGACGTGTGTCAACCA 20653
|||||
Db 47 CTGCTCAGCTCTCTGAGTAGTGGATTACAGGCGCATGCCACCA 1
|||||

RESULT 8
LOCUS AA244245 110 bp mRNA EST 20-AUG-1997
DEFINITION nc07a04.s1 NCI_CGAP.Pr1 Homo sapiens cDNA clone IMAGE:1007406 similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION AA244245
VERSION AA244245.1 GI:1875104
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 110)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 24, 1995 this sequence version replaced gi:634306.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuauqui, M.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -41ml3 fwd. ET from Amersham
High quality sequence stop: 90.

FEATURES
source
1..110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1007406"
/clone_lib="NCI_CGAP_Pr1"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: PAMP10; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected, histologically normal prostate epithelial cells. Double-stranded cDNA was ligated to EcoRI adaptors. 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into PAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

BASE COUNT 17 a 26 c 28 g 38 t 1 others
ORIGIN
Query Match 0.4%; Score 90.4; DB 30; Length 110;
Best Local Similarity 89.0%; Pred. No. 0.6;
Matches 97; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 20507 TTTTGTGAGATGGAGTCTTGTCTGTGCTGCCAGCTGGAGTGGCATGATCTC 20566
|||||
Db 2 TTTTGTGAGATGGAGTCTTGTCTGTGCTGCCAGCTGGAGTGGCATGATCTC 61
|||||

QY 20567 GGTCTACACACCTCCACCTCTGGTTCAGTGATTTCTTGCCTCA 20615
|||||
Db 62 GGCTCACTGCAACCTCTGCTCTCTGGTTCAGAGATTTCTTGCCTCA 110
|||||

RESULT 9
LOCUS AA835205 101 bp mRNA EST 23-FEB-1998
DEFINITION ak64h01.s1 Barstead pancreas HPLRB1 Homo sapiens cDNA clone IMAGE:1412689 3' similar to contains Alu repetitive element;contains element KER repetitive element ;, mRNA sequence.
ACCESSION AA835205
VERSION AA835205.1 GI:2908933
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 101)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Nov 29, 1993 this sequence version replaced gi:636191.
Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

```
FEATURES
  source
Seq primer: -40ml3 fwd. ET from Amersham.
  Location/Qualifiers
    1..101
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:1412689"
    /clone_lib="Barstead pancreas HPLRB1"
    /sex="female"
    /dev_stage="adult, 34 years"
    /lab_host="DH10B"
    /note="Organ: pancreas; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; Site_1: EcoRI; Site_2: NotI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5',
TGTTACGAATCTGAAGTGGGAGCGCGCCCTTTTCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[AAATCGATCGATCTTG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead."
  BASE COUNT      14 a   36 c   27 g   24 t
  ORIGIN

Query Match      0.4%; Score 89.8; DB 39; Length 101;
Best Local Similarity 93.1%; Pred. No. 0.72;
Matches 94; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 14076 TGAGACGAGTCTCACTGTGACCCAGGCTGGAGTGCAGTGGCGGCTCACTCGGCTCACT 14135
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 TGAGACGAGTCTCACTGTGCGCAGGCTGGAGTGCAGTGGCTGATCTCGGCTCACT 60

QY 14136 GCAACCTCGGCTCCGCGTTCAAAGTGAATCTCTCGCTCA 14176
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 GCAAGCTCGGCTCCGCGTTCACGCCATCTCTCGCTCA 101

RESULT 10
AQ264176/c AQ264176 106 bp DNA GSS 27-OCT-1998
LOCUS CITBI-E1-2509A2.TF CITBI-E1 Homo sapiens genomic clone 2509A2,
DEFINITION genomic survey sequence.
ACCESSION AQ264176
VERSION AQ264176.1 GI:3792743
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 106)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT Other GSSs: CITBI-E1-2509A2.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
FEATURES
  source
  Location/Qualifiers
    1..106
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="2509A2"
```

```
/clone_lib="CITBI-E1"
/sex="male"
/cell_type="sperm"
/note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"
  BASE COUNT      25 a   30 c   34 g   17 t
  ORIGIN

Query Match      0.4%; Score 88.4; DB 105; Length 106;
Best Local Similarity 89.6%; Pred. No. 0.99;
Matches 95; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 12739 GGGTTTCACCATGTAGCCAGATGGTCTCGATCTCTGACTCTGATCGCGCCACCTG 12798
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 106 GGGTTTCACCATGTAGCCAGCGGTCTTGATCTCTGACTCTGATCCACCGGCTC 47

QY 12799 AGCTCCCAAGTCTGGGATTACAGGTGTGAGCCACGCGCCCGG 12844
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 46 GGTCTCCCAAGTCTGGGATTACAGCGGTGAGACTCTGCGCCGG 1

RESULT 11
B48914/c B48914 103 bp DNA GSS 08-APR-1999
LOCUS RPC111-4A12.TP RPC111 Homo sapiens genomic clone RPC11-11-4A12,
DEFINITION genomic survey sequence.
ACCESSION B48914
VERSION B48914.1 GI:2601151
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venter,J.C.
TITLE Use of BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPC11-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
FEATURES
  source
  Location/Qualifiers
    1..103
    /organism="Homo sapiens"
    /db_xref="GDB:7501163"
    /db_xref="taxon:9606"
    /clone="RPC11-11-4A12"
    /clone_lib="RPC11-11"
    /sex="Male"
    /cell_type="Lymphocytes"
    /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC library"
  BASE COUNT      30 a   28 c   30 g   15 t
  ORIGIN

Query Match      0.4%; Score 87.8; DB 84; Length 103;
Best Local Similarity 92.9%; Pred. No. 1.2;
Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```

QY 12724 TTTTATAGACAGCGGGTTTACCATTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCG 12783
|||||
Db 99 TTTTATAGACAGCGGGTTTACCCTTTTACCGCGGATGGTCTCGATCTCCTGACCTCG 40
|||||

QY 12784 TGAATCGCCCACTGAGCCTCCCAAGTCTCGGATTAC 12822
|||||
Db 39 TGAATCGCCCACTGAGCCTCCCAAGTCTCGGCTTAC 1
|||||

RESULT 12
AA565533 107 bp mRNA EST 08-SEP-1997
LOCUS n42b11.s1 NCI_CGAP_GC2 Homo sapiens cDNA clone IMAGE:1016157 3'
DEFINITION similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION AA565533
VERSION AA565533.1 GI:2337172
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 107)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1393355.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 1661 Std Error: 0.00
Seq primer: -40ml3 fwd. Et from Amersham
High quality sequence stop: 87.
FEATURES
Location/Qualifiers
1..107
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1016157"
/clone_lib="NCI_CGAP_GC2"
/tissue_type="germ cell tumor"
/lab_host="SOLR (kanamycin resistant)"
/notes="Vector: Bluescript SK-; Site_1: EcoRI; Site_2:
NotI; Cloned unidirectionally. Primer: Oligo dt. Bulk
germ cell tumor. 5' adaptor sequence: 5' GAATTCGGCAGAG 3'
3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTT 3'
Average insert size: 1.2 kb."
BASE COUNT 22 a 34 c 26 g 25 t
ORIGIN

Query Match 0.4%; Score 87.8; DB 35; Length 107;
Best Local Similarity 88.8%; Pred. No. 1.1;
Matches 95; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 14105 CTGGAGTCAGTGGCGGATCTCGGCTCACTGCAAGCTCCGCCCTCCCGGGTTCAAGTGAT 14164
|||||
Db 1 CTGGAGTCAGTGGCGTCACTGAGCTCACTGCAAGCTCTGCTCCAGGTTCAAGTGAT 60
|||||

QY 14165 TCTCTGCTCAGACTCCCGTAGCTGGGATTACAGTCACCA 14211
|||||
Db 61 TCTCTGCTCAGCTCCTCTGAGTAGTGGGATTACAGGCACACCA 107
|||||

```

```

RESULT 13
AA228795 103 bp mRNA EST 20-AUG-1997
LOCUS nc14e07.s1 NCI_CGAP_Pr1 Homo sapiens cDNA clone IMAGE:1008132
DEFINITION similar to contains Alu repetitive element; contains element MER28
repetitive element ;, mRNA sequence.
ACCESSION AA228795
VERSION AA228795.1 GI:1851455
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1394473.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuauqui,
M.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -41ml3 fwd. Et from Amersham
High quality sequence stop: 81.
FEATURES
Location/Qualifiers
1..103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1008132"
/clone_lib="NCI_CGAP_Pr1"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/notes="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was primed with oligo(dt)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected, histologically normal
prostate epithelial cells. Double-stranded cDNA was
ligated to EcoRI adaptors, 5 cycles of PCR applied to the
cDNA with an adaptor-specific primer, and the resulting
PCR product subcloned into pAMP10 by the UDG-cloning
method (Life Technologies). Average insert size is 600
bp. NOTE: Not directionally cloned. This library was
constructed by David Krizman."
BASE COUNT 14 a 30 c 25 g 34 t
ORIGIN

Query Match 0.4%; Score 87; DB 30; Length 103;
Best Local Similarity 90.3%; Pred. No. 1.4;
Matches 93; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3786 TTTTATTTTTTTCAGACGGTGTCTACTCTTCGCCCGCCGCGGACTGCAGTAGCCCTAT 3845
|||||
Db 1 TTTTATTTTTTTCAGATGGTGTCTACTCTGTGCGCCGAGGCTGGAGTGCACTAGCAAT 60
|||||

QY 3846 CTCGGCTCACTGCAAGCTCCGCTCCCGGGTTCACGCCATTTT 3888
|||||
Db 61 CTGGGCTCACTGCAAGCTCCGCTCCCGGGTTCACGCCGTTAT 103
|||||

RESULT 14
AQ582186

```

LOCUS	AQ582186	103 bp	DNA	GSS	07-JUN-1999
DEFINITION	RPCI-11-451A15.TJ RPCI-11 Homo sapiens genomic clone				
ACCESSION	RPCI-11-451A15, genomic survey sequence.				
VERSION	AQ582186				
KEYWORDS	AQ582186.1	GI:5009296			
SOURCE	GSS.				
ORGANISM	human.				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1 (bases 1 to 103)				
JOURNAL	Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.				
COMMENT	Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building				
	Unpublished (1997)				
	On Feb 19, 1999 this sequence version replaced gi:4146076.				
	Other_GSSs: RPCI-11-451A15.TV				
	Contact: Shaying Zhao, William Nierman, Mark Adams				
	Department of Eukaryotic Genomics				
	The Institute for Genomic Research				
	9712 Medical Center Dr., Rockville, MD 20850				
	Tel: 301 838 0200				
	Fax: 301 838 0208				
	Email: hbeetigr.org				
	Clones are derived from the human BAC library RPCI-11. For BAC				
	library availability, please contact Pieter de Jong				
	(pieter@dejong.med.buffalo.edu). Clones may be purchased from				
	BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from				
	Research genet cs (info@resgen.com). BAC end search page:				
	http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.				
	Seq primer: SP6				
	Class: BAC ends.				
FEATURES	Location/Qualifiers				
source	1..103				
	/organism="Homo sapiens"				
	/db_xref="GDB:7672814"				
	/db_xref="taxon:9606"				
	/clone="RPCI-11-451A15"				
	/clone_lib="RPCI-11"				
	/sex="Male"				
	/cell_type="Lymphocytes"				
	/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;				
	RPCI11 Human Male BAC Library"				
BASE COUNT	19 a 36 c 25 g 22 t 1 others				
ORIGIN					
Query Match	0.4%;	Score 86.8;	DB 108;	Length 103;	
Best Local Similarity	91.9%;	Pred. No. 1.5;			
Matches 91;	Conservative	0;	Mismatches 8;	Indels 0;	Gaps 0;
QY 2061	GGAGTCAGTGGCACAATTCGGCTCACTGCACCTCGCCTCCGGGTTCACGCTATTC	20120			
Db 1	GGAGTCGCTGGCACAATTCGGCTCACTGCACCTCGCCTCCAGATTTC	60			
QY 20121	TCTGCTCAGCTCCCAAGTAGCTGGGACTACAGCGC	20159			
Db 61	TCTGCTCAGCTCCGAGTAGCTGGGACTACAGCGC	99			
RESULT 15	AQ584425/c				
LOCUS	AQ584425	103 bp	DNA	GSS	07-JUN-1999
DEFINITION	RPCI-11-458L2.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-458L2				
	genomic survey sequence.				
ACCESSION	AQ584425				
VERSION	AQ584425				
KEYWORDS	GSS.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
	Eutheria; Primates; Catarrhini; Hominidae; Homo.				

Matches 87; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 12715 TTTTGTATTTTGTAGACAGCGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTC 12774
||||| 1 ||||||||||| ||||||||||| || ||||||| || ||||| |||||
Db 1 TTTTGTATCTTTAGTAGACAGAGGGTTTCACCATATGGCCAGGCTGCTCTCAAACTC 60

Qy 12775 CTGACCTCGTATCGGCCACCTGAGCCTCCCAAAGTCTGGGAT 12819

||||| 1 ||||||||||| || ||||||||||| ||||||||||| |||||||
Db 61 CTGACCTTGTGATCACCAGCGCTCGGCCCTCCCAAAGTCTGGGAT 105

RESULT 2

US-08-477-504A-65

; Sequence 65, Application US/08477504A

; Patent No. 5972353

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,504A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3D

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-477-504A-65

Query Match 0.3%; Score 76.2; DB 4; Length 105;

Best Local Similarity 82.9%; Pred. No. 4.9e-06;

Matches 87; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 12715 TTTTGTATTTTGTAGACAGCGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTC 12774

||||| 1 ||||||||||| ||||||||||| || ||||||| || ||||| |||||
Db 1 TTTTGTATCTTTAGTAGACAGAGGGTTTCACCATATGGCCAGGCTGCTCTCAAACTC 60

Qy 12775 CTGACCTCGTATCGGCCACCTGAGCCTCCCAAAGTCTGGGAT 12819

||||| 1 ||||||||||| || ||||||||||| ||||||||||| |||||||
Db 61 CTGACCTTGTGATCACCAGCGCTCGGCCCTCCCAAAGTCTGGGAT 105

RESULT 3

Query Match

Best Local Similarity 82.9%; Score 76.2; DB 4; Length 105;

Matches 87; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 12715 TTTTGTATTTTGTAGACAGCGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTC 12774

||||| 1 ||||||||||| ||||||||||| || ||||||| || ||||| |||||
Db 1 TTTTGTATCTTTAGTAGACAGAGGGTTTCACCATATGGCCAGGCTGCTCTCAAACTC 60

Qy 12775 CTGACCTCGTATCGGCCACCTGAGCCTCCCAAAGTCTGGGAT 12819

||||| 1 ||||||||||| || ||||||||||| ||||||||||| |||||||
Db 61 CTGACCTTGTGATCACCAGCGCTCGGCCCTCCCAAAGTCTGGGAT 105

US-08-486-756A-65

; Sequence 65, Application US/08486756A

; Patent No. 5981711

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,756A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-486-756A-65

Query Match

Best Local Similarity 82.9%; Score 76.2; DB 4; Length 105;

Matches 87; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 12715 TTTTGTATTTTGTAGACAGCGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTC 12774

||||| 1 ||||||||||| ||||||||||| || ||||||| || ||||| |||||
Db 1 TTTTGTATCTTTAGTAGACAGAGGGTTTCACCATATGGCCAGGCTGCTCTCAAACTC 60

Qy 12775 CTGACCTCGTATCGGCCACCTGAGCCTCCCAAAGTCTGGGAT 12819

||||| 1 ||||||||||| || ||||||||||| ||||||||||| |||||||
Db 61 CTGACCTTGTGATCACCAGCGCTCGGCCCTCCCAAAGTCTGGGAT 105

RESULT 4

US-08-485-862B-65

; Sequence 65, Application US/08485862B

; Patent No. 5989838

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court


```

; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIORITY DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-787-739-65

Query Match 0.3%; Score 76.2; DB 5; Length 105;
Best Local Similarity 82.9%; Pred. No. 4.9e-06;
Matches 87; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 12715 TTTTGTGATATTTTAGTAGACAGCGGGTTTCACCATGTTACCCAGGATGGTCTCGATCTC 12774
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 TTTTATCATCTTTAGTAGACAGCGGGTTTCACCATGTTACCCAGGCTCTCAAATC 60

QY 12775 CTGACCTGTGATCGCCACCTGAGCCCTCCAAAGTCTGGGAT 12819
      ||||| ||||| || || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CTGACCTGTGATCCACGAGCTCGGCCCTCCAAAGTCTGGGAT 105

RESULT 6
US-08-454-557C-91
; Sequence 91 Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: Of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:

```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; US-08-454-557C-91

Query Match      0.3%; Score 67; DB 3; Length 84;
Best Local Similarity 88.0%; Pred. No. 0.00023;
Matches 73; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 12747 CCATGTTAGCCAGGATGCTCGATCTCCTGACCTCGTGATCCGCCACCTGAGGCTCCC 12806
      ||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1 CCATGTTTCATCAGGCTGGTGTCGAACCTCCTGACCTCGTGATCCGCCACCTGAGGCTCCC 60
      ||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

QY 12807 AAAGTGCTGGGATTACAGGTGTG 12829
      ||||| ||||| ||||| ||||| |||||
Db 61 AAAGTGCTGGGATTACAGCGGTG 83
      ||||| ||||| ||||| ||||| |||||

RESULT 8
US-08-450-673C-91
; Sequence 91, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; US-08-450-673C-91

Query Match      0.3%; Score 67; DB 4; Length 84;
Best Local Similarity 88.0%; Pred. No. 0.00023;
Matches 73; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 12747 CCATGTTAGCCAGGATGCTCGATCTCCTGACCTCGTGATCCGCCACCTGAGGCTCCC 12806
      ||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1 CCATGTTTCATCAGGCTGGTGTCGAACCTCCTGACCTCGTGATCCGCCACCTGAGGCTCCC 60
      ||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

QY 12807 AAAGTGCTGGGATTACAGGTGTG 12829
      ||||| ||||| ||||| ||||| |||||
Db 61 AAAGTGCTGGGATTACAGCGGTG 83
      ||||| ||||| ||||| ||||| |||||

RESULT 7
US-08-340-426D-91
; Sequence 91, Application US/08340426D
; Patent No. 5948634
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; US-08-340-426D-91

Query Match      0.3%; Score 67; DB 4; Length 84;
Best Local Similarity 88.0%; Pred. No. 0.00023;
Matches 73; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 12747 CCATGTTAGCCAGGATGCTCGATCTCCTGACCTCGTGATCCGCCACCTGAGGCTCCC 12806
      ||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1 CCATGTTTCATCAGGCTGGTGTCGAACCTCCTGACCTCGTGATCCGCCACCTGAGGCTCCC 60
      ||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

QY 12807 AAAGTGCTGGGATTACAGGTGTG 12829
      ||||| ||||| ||||| ||||| |||||
Db 61 AAAGTGCTGGGATTACAGCGGTG 83
      ||||| ||||| ||||| ||||| |||||

RESULT 8
US-08-450-673C-91
; Sequence 91, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; US-08-450-673C-91

Query Match      0.3%; Score 67; DB 4; Length 84;
Best Local Similarity 88.0%; Pred. No. 0.00023;
Matches 73; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 12747 CCATGTTAGCCAGGATGCTCGATCTCCTGACCTCGTGATCCGCCACCTGAGGCTCCC 12806
      ||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1 CCATGTTTCATCAGGCTGGTGTCGAACCTCCTGACCTCGTGATCCGCCACCTGAGGCTCCC 60
      ||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

QY 12807 AAAGTGCTGGGATTACAGGTGTG 12829
      ||||| ||||| ||||| ||||| |||||
Db 61 AAAGTGCTGGGATTACAGCGGTG 83
      ||||| ||||| ||||| ||||| |||||
```


INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-91

Query Match 0.2%; Score 60.6; DB 4; Length 84;
Best Local Similarity 83.1%; Pred. No. 0.0036;
Matches 69; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 7134 CATGCTGTAAATTCAGCACCTTTGGAGCGCGAGCGAGGCAGATCATGAGGTTCAGGAATG 7193

Db 83 CACGCTTGTATCCAGCACCTTTGGAGCGGTGAGCGGCGGATCAGGATCAGGAGTT 24

Qy 7194 CAAGACCAGCCTGACCAATATGG 7216

Db 23 CGACACCAGCCTGATGAACATGG 1

RESULT 12

US-08-450-673C-91/c
Sequence 91, Application US/08450673C
Patent No. 5948888

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/450.673C

FILING DATE: 30-MAY-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609.3840004

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 91:

SEQUENCE CHARACTERISTICS:

LENGTH: 84 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both
US-08-450-673C-91

Query Match 0.2%; Score 60.6; DB 4; Length 84;
Best Local Similarity 83.1%; Pred. No. 0.0036;
Matches 69; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 7134 CATGCTGTAAATTCAGCACCTTTGGAGCGCGAGCGAGGCAGATCATGAGGTTCAGGAATG 7193

Db 83 CACGCTTGTATCCAGCACCTTTGGAGCGGTGAGCGGCGGATCAGGATCAGGAGTT 24

Qy 7194 CAAGACCAGCCTGACCAATATGG 7216

Db 23 CGACACCAGCCTGATGAACATGG 1

RESULT 13

PCT-US95-17111A-91/c

Sequence 91, Application PC/TUS9517111A

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

APPLICANT: Wands, Jack R.

TITLE OF INVENTION: Neural Thread Protein Gene Expression and

TITLE OF INVENTION: Detection of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/17111A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/340,426

FILING DATE: 14-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609.3840002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 91:

SEQUENCE CHARACTERISTICS:

LENGTH: 84 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both
PCT-US95-17111A-91

Query Match 0.2%; Score 60.6; DB 6; Length 84;

Best Local Similarity 83.1%; Pred. No. 0.0036;

Matches 69; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 7134 CATGCTGTAAATTCAGCACCTTTGGAGCGCGAGCGAGGCAGATCATGAGGTTCAGGAATG 7193

Db 83 CACGCTTGTATCCAGCACCTTTGGAGCGGTGAGCGGCGGATCAGGATCAGGAGTT 24

Qy 7194 CAAGACCAGCCTGACCAATATGG 7216

Db 23 CGACACCAGCCTGATGAACATGG 1

RESULT 14

US-08-454-557C-70

Sequence 70, Application US/08454557C

Patent No. 5830670

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

APPLICANT: Wands, Jack R.

TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2000, 12:38:36 ; Search time 17971.8 seconds
(without alignments)
-1569.738 Million cell updates/sec

Title: US-08-852-495C-1_COPY_1_29000
Perfect score: 29000
Sequence: 1 CACACACACACACACACA.....CCAGGCTAGAGTCAGTGGC 29000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_bal.*
- 2: gb_ba2.*
- 3: gb_om.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_pl1.*
- 8: gb_pl2.*
- 9: gb_pr1.*
- 10: gb_pr2.*
- 11: gb_pr3.*
- 12: gb_ro.*
- 13: gb_sts.*
- 14: gb_sy.*
- 15: gb_un.*
- 16: gb_vi.*
- 17: em_fun.*
- 18: em_hum1.*
- 19: em_hum2.*
- 20: em_in.*
- 21: em_om.*
- 22: em_or.*
- 23: em_ov.*
- 24: em_pat.*
- 25: em_ph.*
- 26: em_pl.*
- 27: em_ro.*
- 28: em_sts.*
- 29: em_sy.*
- 30: em_un.*
- 31: em_vi.*
- 32: gb_htg1.*
- 33: gb_htg2.*
- 34: gb_in1.*
- 35: gb_in2.*
- 36: em_bal.*
- 37: em_ba2.*
- 38: em_hum3.*
- 39: em_hum4.*
- 40: gb_pr4.*
- 41: gb_htg3.*
- 42: gb_htg4.*
- 43: gb_htg5.*
- 44: gb_htg6.*

- 45: gb_htg7.*
- 46: em_htg1.*
- 47: em_htg2.*
- 48: em_htg3.*
- 49: em_hum5.*
- 50: gb_pl3.*
- 51: gb_pr5.*
- 52: gb_htg8.*
- 53: gb_htg9.*
- 54: gb_htg10.*
- 55: gb_htg11.*
- 56: gb_htg12.*
- 57: gb_htg13.*
- 58: gb_htg14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
C 1	87.2	0.3	103	9 HUMALCE221	M87896 Human carci
C 2	87	0.3	107	9 HUMALCE162	M87924 Human carci
C 3	87.2	0.3	108	10 HSLDLRN2	X05250 Human LDL-r
C 4	83.6	0.3	108	10 HSLDLRN2	X05250 Human LDL-r
C 5	79.8	0.3	108	10 HSLDLRD1	X05249 Human LDL-r
C 6	79.8	0.3	108	10 HSLDLRD2	X05251 Human LDL-r
C 7	80	0.3	108	11 HSU67803	U67803 Human small
C 8	75	0.3	103	9 HUMALCE221	M87896 Human carci
C 9	75.2	0.3	108	11 HSU67804	U67804 Human small
C 10	74.2	0.3	108	9 HUMDL1003M5	D16965 Human HepG2
C 11	73.6	0.3	108	10 HSLDLI12	X05248 Human LDL-r
C 12	73.4	0.3	110	11 HSU67807	U67807 Human small
C 13	73	0.3	103	13 HS8IC8R	X57789 Human sequ
C 14	73	0.3	108	10 HSLDLRD1	X05249 Human LDL-r
C 15	73	0.3	108	10 HSLDLRD2	X05251 Human LDL-r
C 16	72.4	0.2	101	10 S79560	S79560 HRX {intron
C 17	71.6	0.2	94	9 HUMHGAL	M13479 Human alpha
C 18	70.8	0.2	90	9 HUMDLRFL	K03555 Human low d
C 19	70.8	0.2	91	13 HUMUT8164A	L30244 Human STS U
C 20	69.8	0.2	108	13 G32614	G32614 A009K21 Hum
C 21	69.8	0.2	110	9 HUMALCE43	M87900 Human carci
C 22	68.8	0.2	106	13 G32743	G32743 A009P31 Hum
C 23	69	0.2	108	11 HSU67803	U67803 Human small
C 24	68.8	0.2	108	11 HSU67808	U67808 Human small
C 25	68.4	0.2	95	13 HUMUT8002B	L30176 Human STS U
C 26	68	0.2	97	9 HUMDLRLA1	M14178 Human low d
C 27	67.8	0.2	100	9 HUMGALNSA	D45223 Human GALNS
C 28	68	0.2	103	13 HS8IC8R	X57789 Human sequ
C 29	68	0.2	104	9 HUMALCE272	M87899 Human carci
C 30	68	0.2	107	9 HUMALCE162	M87924 Human carci
C 31	67.8	0.2	108	13 G43535	G43535 WIAF-2393-S
C 32	67.2	0.2	85	10 HUMHIS1PR	M26162 Homo sapien
C 33	67.2	0.2	97	9 HUMDLRA2	M14180 Human low d
C 34	66.4	0.2	97	9 HUMDLRA2	M14180 Human low d
C 35	65.8	0.2	99	13 HUMUT7692A	L30306 Human STS U
C 36	66	0.2	100	13 G43536	G43536 WIAF-2394-S
C 37	66	0.2	100	13 G43538	G43538 WIAF-2396-S
C 38	66	0.2	110	9 HUMALCE43	M87900 Human carci
C 39	65.6	0.2	80	9 HUMBRKFAE	M36135 Human alpha
C 40	65.6	0.2	107	11 HSU67806	U67806 Human small
C 41	65.2	0.2	79	10 S73203	S73203 ALL-1 {tand
C 42	65.4	0.2	110	11 HSU67807	U67807 Human small
C 43	64.6	0.2	95	10 HSSTHPKIB	X66361 H.sapiens m
C 44	64.8	0.2	96	4 NVIHIS2A	J00950 Newt histon
C 45	64.4	0.2	97	9 HUMDLRDJ	M14179 Human faml1

ALIGNMENTS

```

RESULT 1
HUMALCE221/c
LOCUS HUMALCE221 103 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
Location/Qualifiers
source
1..103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Ntera2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT 25 a 27 c 33 g 18 t
ORIGIN
Query Match 0.3%; Score 87.2; DB 9; Length 103;
Best Local Similarity 92.0%; Pred. No. 9.1e-06;
Matches 92; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 12205 CTTGAGTCAATGGCGGCGATCTTGCTCACAGCAACCTCCGCTCCGGTTCAAGCCAT 12264
|||||
Db 103 CTTGAGTCAATGGCGGCGATCTTGCTCACAGCAACCTCCGCTCCGGTTCAAGCCAT 44
|||||

QY 12265 TCTCTGCTCAGCTCCGCTGAGTCTGGGATTACAGGCA 12304
|||||
Db 43 TCTCTGCTCAGCTCCGCTGAGTCTGGGATTACAGGCA 4
|||||

RESULT 2
HUMALCE162/c
LOCUS HUMALCE162 107 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE162.
ACCESSION M87924
VERSION M87924.1 GI:174871
KEYWORDS Alu repeat.
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 107)
AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
Location/Qualifiers
source
1..107
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Ntera2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT 28 a 30 c 35 g 14 t
ORIGIN
Query Match 0.3%; Score 87; DB 9; Length 107;
Best Local Similarity 90.3%; Pred. No. 9.8e-06;
Matches 93; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

```

QY 3564 TTTTGTGAGAGGAGCTAGTCTGTGCGCCAGGCTGGAGTGGCAGCCATCTTGGC 3623
|||||
Db 107 TTTTGTGAGAGGAGCTAGTCTGTGCGCCAGGCTGGAGTGGCAGCCATCTTGGC 48
|||||

QY 3624 TCATGCAAGCTCTGCCCTCCCGGGTTATGCCATTTCTCATGTC 3666
|||||
Db 47 TCATGCAAGCTCTGCCCTCCCGGGTTATGCCATTTCTTCTGCC 5
|||||

RESULT 3
HSLDLRN2/c
LOCUS HSLDLRN2 108 bp DNA PRI 20-MAY-1992
DEFINITION Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION X05250
VERSION X05250.1 GI:34337
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
Williamson,R. and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901
COMMENT See X05252 for deletion junction
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES
Location/Qualifiers
source
1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..108
/note="intron XIV fragment"
BASE COUNT 28 a 23 c 39 g 18 t
ORIGIN
Query Match 0.3%; Score 87.2; DB 10; Length 108;
Best Local Similarity 88.0%; Pred. No. 9.1e-06;
Matches 95; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 3618 CTTGGCTCACTGCACAGCTCTGCCTCCCGGGTTATGCCATTTCTCATGTCACGCTCCAG 3677
|||||
Db 108 CTTGGCTCACTGCACAGCTCTGCCTCCCGGGTTATGCCATTTCTCATGTCACGCTCCAG 49
|||||

QY 3678 AGTAGCTGGGACTACAGCGCCGCCACACACGCTGGCTAATTTTTT 3725
|||||
Db 48 AGTAGCTGGGATTACAGGCACCTGCCACCACGCTGGCTAATTTTTGT 1
|||||

RESULT 4
HSLDLRN2
LOCUS HSLDLRN2 108 bp DNA PRI 20-MAY-1992
DEFINITION Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION X05250
VERSION X05250.1 GI:34337
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
Williamson,R. and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901

```


COMMENT See X05252 for deletion junction
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.

FEATURES

source
1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..108
/note="intron XIV fragment"
BASE COUNT 28 a 23 c 39 g 18 t
ORIGIN

Query Match 0.3%; Score 83.6; DB 10; Length 108;
Best Local Similarity 86.8%; Pred. No. 3.8e-05;
Matches 92; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 24520 AAAAAATGACGAGCATGTGTGGCGGCTATATCCAGCTAATTTGGGAGGCTGAGGC 24579
|||||
Db 3 AAAAAATGACGAGCGTGTGGCAGGTGCTGTATCCAGCTACTCGGGAGGCTGAGGC 62
QY 24580 AGGAGAAATGCCTGACCTGGGAGGTGGAGTTGCACTGAGCCAAAG 24625
|||||
Db 63 AGGAGAAATGCTTGAACCCAGGAGGAGGTTGCACTGAGCCGAG 108

RESULT 5

LOCUS HSLDLRD1 108 bp DNA PRI 20-MAY-1992
DEFINITION Human LDL-receptor mutated gene with intron 12 deletion junction.
ACCESSION X05249
VERSION X05249.1 GI:34335
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
Williamson,R. and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901
COMMENT *source: hypercholesterol aemia
See X05248 for corresponding normal gene sequence
In the defective LDL-receptor gene the deletion occurred between two alu-repetitive sequences, that are in the same direction, the deletion eliminates exons 13 and 14 and changes the reading frame of the resulting spliced mRNA.
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.

FEATURES

source
1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="blood leukocytes from a patient with familial"
misc_feature 1..108
/note="deletion junction region intron 12/ intron 15"
BASE COUNT 20 a 40 c 20 g 28 t
ORIGIN

Query Match 0.3%; Score 79.8; DB 10; Length 108;
Best Local Similarity 84.1%; Pred. No. 0.00018;
Matches 90; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 3619 TTGGCTCACTGCAAGCTCTGCCCTCCGGGTTATGCAATCTCATGCTCAGCCTCAGA 3678
|||||
Db 2 TCGCCTCACCACAACTCTGCTCTCTGGGTTCAACCAATTTCTGCTCAGCCTCCGA 61
QY 3679 GTAGCTGGGACTACAGCGGCCGCCACCGCTGGCTAATTTTTT 3725
|||||
Db 62 GTAGCTGGGATTACAGGCACCTGCCACCGCTGGCTAATTTTGT 108

RESULT 6

LOCUS HSLDLRD2/c 108 bp DNA PRI 20-MAY-1992
DEFINITION Human LDL-receptor mutated gene with intron 14 deletion junction.
ACCESSION X05251
VERSION X05251.1 GI:34336
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
Williamson,R. and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901
COMMENT *source: hypercholesterol aemia
See X05250 for corresponding normal gene sequence
In the defective LDL-receptor gene the deletion occurred between two alu-repetitive sequences, that are in the same direction, the deletion eliminates exons 13 and 14 and changes the reading frame of the resulting spliced mRNA.
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.

FEATURES

source
1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="blood leukocytes from a patient with familial"
intron 1..108
/note="intron XIV fragment"
BASE COUNT 28 a 20 c 40 g 20 t
ORIGIN

Query Match 0.3%; Score 79.8; DB 10; Length 108;
Best Local Similarity 84.1%; Pred. No. 0.00018;
Matches 90; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 3619 TTGGCTCACTGCAAGCTCTGCCCTCCGGGTTATGCAATCTCATGCTCAGCCTCAGA 3678
|||||
Db 107 TCGCCTCACCACAACTCTGCTCTCTGGGTTCAACCAATTTCTGCTCAGCCTCCGA 48
QY 3679 GTAGCTGGGACTACAGCGGCCGCCACCGCTGGCTAATTTTTT 3725
|||||
Db 47 GTAGCTGGGATTACAGGCACCTGCCACCGCTGGCTAATTTTGT 1

RESULT 7

LOCUS HSU67803/c 108 bp RNA PRI 01-AUG-1997
DEFINITION Human small cytoplasmic Alu transcript.
ACCESSION U67803
VERSION U67803.1 GI:2289917
KEYWORDS Alu.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE cDNAs derived from primary and small cytoplasmic Alu (scAlu) transcripts
JOURNAL J. Mol. Biol. 271 (2), 222-234 (1997)
MEDLINE 97415756
REFERENCE 2 (bases 1 to 108)
AUTHORS Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The

Children's Hospital of Philadelphia, 1004F Abramson Research
Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

```
FEATURES
  source
    1..108
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="TscAlu2"
    1..108
      /note="scAlu"
      /rpt_family="Alu"
      /rpt_type="dispersed"
  repeat_region
    23 a 39 c 30 g 16 t

BASE COUNT
ORIGIN
  23 a 39 c 30 g 16 t

Query Match
  0.3%; Score 80; DB 11; Length 108;
Best Local Similarity 89.6%; Pred. No. 0.00016;
Matches 86; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3735 GTAGAGATGGGTTTCCACCGTGTAGCCAGAAATGCTCGATCTCTTGACCTTCGTGATCC 3794
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 97 GTAGAGACGGGTTTCACCTTGTAGCCAGGATGCTCGATCTCTTGACCTTCGTGATCC 38
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 3795 GCCTGCTTGGTTCCTCCAAAGTCTGGGATACAG 3830
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 37 GCCCGCTCGGCTCCCAAGTCTGGGATACAG 2
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
HUMALCE221 HUMALCE221 103 bp ss-RNA PRI 15-APR-1994
LOCUS Human carcinoma cell-derived Alu RNA transcript, clone CE221.
DEFINITION M87896
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of
  post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
  Location/Qualifiers
    source
      1..103
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /cell_line="NTERa2D1"
        /dev_stage="embryo"
        /sex="male"
        /tissue_type="carcinoma"
  BASE COUNT
  ORIGIN
    25 a 27 c 33 g 18 t

Query Match
  0.3%; Score 75; DB 9; Length 103;
Best Local Similarity 84.8%; Pred. No. 0.0012;
Matches 84; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 24547 GCCTATATCCAGCTAATTTGGAGGCTGAGCGAGGAGAAATTCCTGAACTGGGAGGTG 24606
||||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5 GCCTGTATATCCAGCTACACGGGAAGTAAGCGAGGAGAAATTCCTGAACTGGGAGGCG 64
||||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 24607 GAGTGTGACTGAGCAGATCACACCATTCGATCCAG 24645
||||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 65 GAGTGTGAGTGGCGGAGATCGTGCCATTCGATCCAG 103
||||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
HSU67804/c HSU67804 108 bp RNA PRI 01-AUG-1997
LOCUS Human small cytoplasmic Alu transcript.
DEFINITION
```

```
ACCESSION U67804
VERSION U67804.1 GI:2289918
KEYWORDS Alu.
SOURCE human.
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
    Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1 (bases 1 to 108)
  AUTHORS Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE cDNAs derived from primary and small cytoplasmic Alu (scAlu)
    transcripts
  JOURNAL J. Mol. Biol. 271 (2), 222-234 (1997)
  MEDLINE 97415756
  REFERENCE 2 (bases 1 to 108)
  AUTHORS Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE Direct Submission
  JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
    Children's Hospital of Philadelphia, 1004F Abramson Research
    Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
  FEATURES
    Location/Qualifiers
      source
        1..108
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
          /clone="TscAlu3"
        1..108
          /note="scAlu"
          /rpt_family="Alu"
          /rpt_type="dispersed"
      repeat_region
        26 a 38 c 26 g 18 t

BASE COUNT
ORIGIN
  26 a 38 c 26 g 18 t

Query Match
  0.3%; Score 75.2; DB 11; Length 108;
Best Local Similarity 86.5%; Pred. No. 0.0011;
Matches 83; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 3735 GTAGAGATGGGTTTCCACCGTGTAGCCAGAAATGCTCGATCTCTTGACCTTCGTGATCC 3794
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 97 GGAAGACGGGTTTCCACCATGTTAGCCAGGATGCTCGATCTCTTGACCTTCGTGATCC 38
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 3795 GCCTGCTTGGTTCCTCCAAAGTCTGGGATACAG 3830
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 37 TCCCGCTTGGCTTCCAAAGTCTGGGATACAG 2
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
HUMD1D03M5/c HUMD1D03M5 108 bp mRNA PRI 04-FEB-1999
LOCUS Human HepG2 partial cDNA, clone hmdid03m5.
DEFINITION D16965
ACCESSION D16965
VERSION D16965.1 GI:598552
KEYWORDS gene signature.
SOURCE Homo sapiens Male cell_line:HepG2 cDNA to mRNA, clone_lib:Kiseru.
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
    Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1 (bases 1 to 108)
  AUTHORS Matoba,R.
  TITLE Direct Submission
  JOURNAL Submitted (21-JUL-1993) to the DBJ/EMBL/GenBank databases. Ryo
    Matoba, Osaka University, Institute for Molecular and Cellular Bio;
    1-3, Yamada-oka, Suita, Osaka 565, Japan
    (E-mail:matoba@inherit.imcb.osaka-u.ac.jp,
    Tel:81-6-877-5111(ex.3314), Fax:81-6-877-1922)
  REFERENCE 2 (bases 1 to 108)
  AUTHORS Matoba,R., Okubo,K., Hori,N., Fukushima,A. and Matsubara,K.
  TITLE The addition of 5'-coding information to a 3'-directed cDNA library
    improves analysis of gene expression
  JOURNAL Gene 146 (2), 199-207 (1994)
  MEDLINE 94357437
  COMMENT Submitted (21-Jul-1993) to DBJ by:
    Ryo Matoba
    Molecular Microbiology and Genetics Lab.
```

Research Institute of Innovative Technology for the Earth 9-2
Kizugawadai Kizu-cyo,
Soraku-gun, Kyoto
Japan, 619-02
Phone: 07747-5-2308
Fax: 07747-5-2321.

FEATURES

source
1. .108
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Hep62"
/clone_lib="Kiseru"
/sex="Male"
BASE COUNT 28 a 23 c 38 g 17 t 2 others
ORIGIN

Query Match 0.3%; Score 74.2; DB 9; Length 108;
Best Local Similarity 85.3%; Pred. No. 0.0017;
Matches 93; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 12222 GATCTGGCTACAGCAACCTCCGGCTCCGGGTTCAAGCCATTCTCGCTCAGCCCTC 12281
|||||
Db 108 GATCTGGCTACATGCAACTCTGCTCCGGGTTCAAGGACTCTCTCGCTCAGCCCTC 49
|||||

QY 12282 CGGAGTAGCTGGGATTACAGGATCGCCACGACACCTCGCTGCTAAATTTT 12330
|||||

Db 48 CTGAGTAGCTGGGATTACAGGATCGCCACGACACCTCGCTGCTTTTAT 1
|||||

RESULT 11

HSDDL112 108 bp DNA PRI 20-MAY-1992
LOCUS
DEFINITION Human LDL-receptor gene intron 12 fragment (normal gene) LDL = low density lipoprotein.

ACCESSION X05248

VERSION X05248.1 GI:34334

KEYWORDS Alu repetitive sequence; low density lipoprotein receptor; repetitive sequence.

SOURCE

human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
Williamson,R. and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia

JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)

MEDLINE 87161901

COMMENT see X05249 for deletion junction

Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.

FEATURES

source
1. .108
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature complement(<1..65)
/note="Alu repeat"
intron 1. .108
/note="intron XII fragment"

BASE COUNT 21 a 38 c 20 g 29 t

ORIGIN

Query Match 0.3%; Score 73.6; DB 10; Length 108;
Best Local Similarity 81.7%; Pred. No. 0.0021;
Matches 85; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 12226 TTGGCTACAGCAACCTCCGGGTTCAAGGATTCTCTCGCTCAGCCCTCGGA 12285
|||||

Db 2 TCGCCTCACCACAACCTCTGCCCTCGGGTTCAAGGATTCTCTCGCTCAGCCCTCCTTA 61
|||||

QY 12286 GTAGCTGGGATTACAGGATCGCCACGACACCTCGCTAAATTT 12329
|||||

Db 62 GTAGCTGGGATTACAGGATGTGCCACCGCCCGCTGATTTT 105
|||||

RESULT 12

HSU67807/c 110 bp RNA PRI 01-AUG-1997
LOCUS
DEFINITION Human small cytoplasmic Alu transcript.

ACCESSION U67807

VERSION U67807.1 GI:2289921

KEYWORDS Alu.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 110)

AUTHORS

Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.

TITLE

cDNAs derived from primary and small cytoplasmic Alu (scAlu) transcripts

JOURNAL J. Mol. Biol. 271 (2), 222-234 (1997)

MEDLINE 97415756

REFERENCE 2 (bases 1 to 110)

AUTHORS Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.

TITLE Direct Submission

JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The Children's Hospital of Philadelphia, 1004F Abramson Research Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

FEATURES

Location/Qualifiers
1. .110

source /organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="TscAlu6"

repeat_region 1. .110

/note="scAlu"

/rpt_family="Alu"

/rpt_type="dispersed"

BASE COUNT 26 a 39 c 24 g 21 t

ORIGIN

Query Match 0.3%; Score 73.4; DB 11; Length 110;
Best Local Similarity 83.8%; Pred. No. 0.0023;
Matches 83; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 2864 GTAGAGATGGGGTTTCACTATGTGGCCAGGCTAGTTTGAACCTCTGACCTCAGTGAT 2923
|||||

Db 99 GGAAGATGGGGTTTCACTATGTGGCCAGGCTAGTTTGAACCTCTGACCTCAGTGAT 40
|||||

QY 2924 CCATTCTCATTTGGCTCCCAAGTGTGGGATTACAGGC 2962
|||||

Db 39 CCACCCACATTTGGCCCTCTCAAGTGTGGGATTACAGGC 1
|||||

RESULT 13

HS81C8R 103 bp DNA STS 05-SEP-1991
LOCUS
DEFINITION Human sequence tagged site 81C8R DNA from 19q13.

ACCESSION X57789

VERSION X57789.1 GI:23938

KEYWORDS STS; myotonic dystrophy.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 103)

AUTHORS Aldridge,F.L.

TITLE Direct Submission

JOURNAL Submitted (12-FEB-1991) F.L. Aldridge, ICI Pharmaceuticals,

Allderley Park, Macclesfield, Cheshire, SK10 4TG, UK

REFERENCE 2 (bases 1 to 103)

AUTHORS Butler,R., Riley,J.H., Ogilvie,D.J., Anand,R., Buxton,J.,
Davies,J., Johnson,K. and Markham,A.F.

RESULT	15
LOCUS	HSLDLRD2
DEFINITION	Human LDL-receptor mutated gene with intron 14 deletion junction
ACCESSION	X05251
VERSION	X05251.1
KEYWORDS	Alu repetitive sequence; low density lipoprotein receptor.
SOURCE	human.

Eukaryota, Metazoa, Chordata, Vertebrata, Mammalia, Eutheria,
 Primates, Catarrhini; Hominidae; Homo.
 1 (bases 1 to 108)
 AUTHORS
 Horsthemke, B., Beisiegel, U., Dunning, A., Havinga, J.-R.
 REFERENCE

TITLE Unequal crossing-over between two alu-repetitive DNA sequences in Williamson, K. and Humphries, S.

JOURNAL MEDLINE
EUI: J. BIOCHEM. 104 (1), 77-81 (1987)
87161901
*source: hypercholesterol aemia
See X05250 for corresponding normal gene sequence
COMMENT

```

FEATURES
source
1..108
/organism="Homo sapiens"
/db_xref="taxon.9606"
/cell_type="blood leukocytes from a patient with familial"
1..108
intron

```

BASE COUNTY 28 a 40 g 20 L
ORIGIN

Query Match 0.3%; Score 73; DB 10; Length 108;
Best Local Similarity 81.0%; Pred. No. 0.0027;
Matches 85; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Db	3	AAAAATTAGCAGCGTGGTGGAGGTGCCTGTAATCCCACTACTCCGGAGGCTGAGGC	62
QY	24580	AGGAGAAATTCCTCAACTGGGAGGTGGAGGTTGCACTGAGCCAA	24624
Db	63	AGGAAATGTTTAAACCCAGGAGGCAGAGGTTCTGTGTAGGCCA	107

Search completed: June 14, 2000, 20:22:29
Job time: 27833 sec

Search completed: June 14, 2000, 20:22:29
Job time: 27833 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2000, 12:58:09 ; Search time 593.67 Seconds
(without alignments)
12221.555 Million cell updates/sec

Title: US-08-852-495C-1_COPY_1_29000
 Perfect score: 29000
 Sequence: 1 CACACACACACACACACA.....CCAGGCTAGAGTGCAGTGGC 29000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 433070

```
Minimum DB seq length: 10
Maximum DB seq length: 110
```

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:★

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	C	68.4	0.2	108	1	X12095	Human bla1lelic po
2	C	66.6	0.2	100	1	X12085	Human bla1lelic po
3	C	66.6	0.2	100	1	X12086	Human bla1lelic po
4	C	65.4	0.2	86	1	V41231	Mouse embryonic ce
5	C	65	0.2	100	1	X12087	Human bla1lelic po
6	C	63.4	0.2	100	1	X12087	Human bla1lelic po
7	C	63.4	0.2	100	1	X12085	Human bla1lelic po
8	C	63.4	0.2	100	1	X12086	Human bla1lelic po
9	C	62.8	0.2	100	1	Q76490	Human genome fragm
10	C	62.2	0.2	108	1	X12095	Human bla1lelic po
11	C	60.8	0.2	100	1	T24892	Human gene signatu
12	C	59.2	0.2	102	1	T20743	Human gene signatu
13	C	58.6	0.2	103	1	T20927	Human gene signatu
14	C	57.8	0.2	103	1	T26213	Human gene signatu
15	C	56.4	0.2	84	1	T25848	Human gene signatu
16	C	56.2	0.2	108	1	T25009	Human gene signatu
17	C	56.2	0.2	108	1	T26828	Human gene signatu
18	C	55.4	0.2	110	1	T25260	Human gene signatu
19	C	54.8	0.2	87	1	T23566	Human gene signatu
20	C	53.6	0.2	69	1	Q29016	Probe to internal
21	C	53	0.2	106	1	Q95210	Simple tandem repe
22	C	52.6	0.2	110	1	T25260	Human gene signatu
23	C	52.2	0.2	65	1	T25588	Human gene signatu
24	C	52	0.2	93	1	T25688	Human gene signatu
25	C	52.2	0.2	103	1	T20927	Human gene signatu
26	C	52	0.2	108	1	T26828	Human gene signatu
27	C	51.4	0.2	70	1	N60231	Normal chromosome
28	C	51.6	0.2	110	1	T26388	Human gene signatu
29	C	51.2	0.2	100	1	T24892	Human gene signatu
30	C	50.6	0.2	99	1	T23728	Human gene signatu
31	C	49.8	0.2	91	1	T25854	Human gene signatu
32	C	50	0.2	99	1	T20931	Human gene signatu
33	C	49.6	0.2	93	1	T23572	Human gene signatu
34	C	49.2	0.2	108	1	T25009	Human gene signatu

ALIGNMENTS

RESULT	1	
X12095		
ID	X12095 standard; DNA; 108 BP.	
AC	X12095;	
DT	30-MAR-1999 (first entry)	
DE	Human biallelic polymorphic DNA fragment TIGR-A003M18a.	
KW	Polymorphism; biallelic; human; forensic; paternity testing; disease;	
KW	detection; phenotypic typing; characteristic; infection; hereditary;	
KW	autoimmune disease; cancer; inflammation; drug; therapy; medicament;	
KW	treatment; marker; ss.	
OS	Homo sapiens.	
PN	W09820165-A2.	
PD	14-MAY-1998.	
PF	05-NOV-1997; U20313.	
PR	05-NOV-1996; US-030455.	
PA	(WHED) WHITEHEAD INST BIOMEDICAL RES.	
PI	Hudson T, Lander ES, Wang D;	
PT	WPI; 98-286374/25.	
DR	New isolated nucleic acid segments from the human genome - used for	
PT	determining polymorphic forms for use in e.g. forensics, paternity	
PT	testing or phenotypic typing for disease	
PS	Claim 1; Page 219; 31opp; English.	
CC	X10269-X12937 are human DNA fragments which contain biallelic polymorphic	
CC	markers which have been isolated using the primers represented in	
CC	X09121-X10268. The base occupying the polymorphic site is indicated by	
CC	the appropriate IUPAC-RUB ambiguity code. These fragments can be used in	
CC	methods for determining polymorphic forms in an individual for use in	
CC	e.g. forensics, paternity testing or for phenotypic typing for diseases	
CC	such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,	
CC	muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial	
CC	hypercholesterolemia, polycystic kidney disease, hereditary	
CC	spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary	
CC	haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos	
CC	syndrome, osteogenesis imperfecta, acute intermittent porphyria,	
CC	autoimmune diseases, inflammation, cancer, diseases of the nervous	
CC	system, infection by pathogenic microorganisms, and characteristics such	
CC	as longevity, appearance (e.g. baldness, obesity), strength, speed,	
CC	endurance, fertility, and susceptibility or receptivity to particular	
CC	drugs or therapeutic treatments. The isolated polymorphic nucleic acid	
CC	segments can also be used to produce medicaments for the treatment or	
CC	prophylaxis of such diseases.	
Sequence	108 BP; 19 A;	23 C; 28 G; 37 T;

	Query Match	0.28;	Score 68.4;	DB 1;	Length 108;
	Best Local Similarity	82.4%;	Pred. No. 0.033;		
	Matches	89;	Conservative	1;	Mismatches 17;
					Indels 1;
					Gaps 1;
Qy	2854	TGTAATTTT	TAGTAGAGATGGGGTTT	TCAC	TATGTGGCCAGGCTAGTTTGGAACTCCTGAC
Db	1	TGTC	TTTTTTGTAGAGATGAGGTTT	CTC	TRTTGGCCAGGATGCTCGAACTCCTGAC
					60
Qy	2914	CTC	CAGTGA	CCATTC	TCATTTGGGCTCCC-AAAGTCTGGGATACAG
Db	61	TTCAAGT	GATCCG	TCCTGC	TTGGGCTCCCAAAGTCTGGGATATAG
					108

```
RESULT 2
X12085/C
ID X12085 standard; DNA; 100 BP.
AC X12085;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment EST98276c.
KW polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medication;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
FA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 218; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-TUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 22 A; 25 C; 22 G; 30 T;

Query Match 0.28; Score 66.6; DB 1; Length 100;
Best Local Similarity 78.88; Pred. No. 0.061;
Matches 78; Conservative 1; Mismatches 20; Indels 0; Gaps 0;

QY 23012 GTGGCTCATGCTCTAATCCAGACACTTTGAGAGGCTGAAGAGGAGGATCGTTGAGTC 23071
   ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 99 GTGACTCACACCTATAATCTGGCAGCTTTAGGAGGCTTAGGAGGAGGATTTTGAAC 40

QY 23072 CGGAGTTCAGAGACCTCTGGGCAACACAGCAGGACCC 23110
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 39 CAGGAGCTCAAGACCATCTCTGGGAACATAGCAAGATC 1

RESULT 4
V41231/C
ID V41231 standard; cDNA; 86 BP.
AC V41231;
DT 01-OCT-1998 (first entry)
DE Mouse embryonic cell EST 13-4 nucleotide sequence.
KW Embryonic stem cell; ESC; non-primate; mouse; EST; human;
KW developmental gene; transgenic animal; reporter gene; ss.
OS Mus sp.
PN W09823633-A1.
PD 04-JUN-1998.
PF 25-NOV-1997; U22335.
PR 27-NOV-1996; US-032510.
PA (CORR ) CORNELL RES FOUND INC.
PI Holtschu DL, Mark WH;
DR WPI; 98-322656/28.
PT Screening for human developmental genes - by trapping in murine
PT embryonic stem cells and analysing differential expression in vitro,
PT selecting homologous non-human primate gene and using it to isolate
PT human gene
PS Claim 37; Page 18; 60pp; English.
CC Sequences shown in V41230 to V41247 represent nucleotide sequences of
CC mouse EST from tagged cDNA clones. These are used in the method of the
CC invention of screening for human developmental genes. The method
CC comprises inserting a promoterless reporter gene into a non-primate
CC mammalian embryonic stem cell (ESC) genome and identifying cellular
CC transcripts that encode the reporter gene product. Fragments of genes
CC encoding these transcripts are cloned and sequenced. A gene encoding a
CC transcript that includes unknown sequences is selected and expression
CC level of the gene encoding the transcript, or part of it, in different
CC cell types and/or different developmental stages is detected. A gene
CC showing differential expression is selected and expression levels of a
```


CC homologous non-human primate gene, in different cell types and/or at
 CC different developmental stages, using the non-primate transcript as
 CC probe is detected. A homologous gene having the same pattern of
 CC differential expression is selected and the non-primate gene, or part of
 CC it is used to identify the homologous human gene. The ESC transcripts
 CC identified by this method are used to generate transgenic animals
 CC baboon or chimpanzee for study of gene function. The method provides
 CC rapid and large scale screening for human developmental genes, and
 CC eliminates the need to analyse reporter gene expression in embryos.
 SQ Sequence 86 BP; 16 A; 28 C; 30 G; 12 T;

Query Match 0.2%; Score 65.4; DB 1; Length 86;
 Best Local Similarity 86.7%; Pred. No. 0.09;
 Matches 72; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 9626 TGACTCTCGTCTTCTTAGCGAGACGCGCTGGATTTAGGAGGACGCGCCCTGA 9685

DB 84 TGGCTCTCGTCTTCTTGGCGACGACGCGCTGGATTTGGCAGGACGCGCCCTGC 25

QY 9686 GCAATGGTCACCGGCTAGCAG 9708

DB 24 GCGATGGTCACGCGCCACGAG 2

RESULT 5

X12087/C

ID X12087 standard; DNA; 100 BP.

AC X12087;

DT 30-MAR-1999 (first entry)

DE Human biallelic polymorphic DNA fragment EST98276a.

KW Polymorphism; biallelic; human; forensic; paternity testing; disease;

KW detection; phenotypic typing; characteristic; infection; hereditary;

KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;

KW treatment; marker; ss.

OS Homo sapiens.

PN W09820165-A2.

PD 14-MAY-1998.

PF 05-NOV-1997; U20313.

PR 06-NOV-1996; US-030455.

PA (WHED) WHITEHEAD INST BIOMEDICAL RES.

PI Hudson T, Lander ES, Wang D;

DR WPI; 98-286974/25.

PT New isolated nucleic acid segments from the human genome - used for

PT determining polymorphic forms for use in e.g. forensics, paternity

PT testing or phenotypic typing for disease

PS Claim 1; Page 219; 310pp; English.

CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic

CC markers which have been isolated using the primers represented in

CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in

CC methods for determining polymorphic forms in an individual for use in

CC e.g. forensics, paternity testing or for phenotypic typing for diseases

CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,

CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial

CC hypercholesterolemia, polycystic kidney disease, hereditary

CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary

CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos

CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,

CC autoimmune diseases, inflammation, cancer, diseases of the nervous

CC system, infection by pathogenic microorganisms, and characteristics such

CC as longevity, appearance (e.g. baldness, obesity), strength, speed,

CC endurance, fertility, and susceptibility or receptivity to particular

CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid

CC segments can also be used to produce medicaments for the treatment or

CC prophylaxis of such diseases.

SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match

Best Local Similarity 0.2%; Score 65; DB 1; Length 100;

Matches 77; Conservative 1; Mismatches 21; Indels 0; Gaps 0;

QY 23012 GTGGCTCATGCTTAATCCAGCACTTTGAGAGGCTGAAGAGGAGGATCGTTGAGTC 23071

DB 99 GTGACTCACCTATATATCTTGGCACTTTAGGAGGCTTAGGAGGAGGATGTTTGAAC 40

QY 23072 CGGGAGTTTCAAGGATCTCTGGGCAACACAGCGAGACCC 23110

DB 39 CAGGAGCTCAAGACCAKCTGTGGNAACATAGCAAGACTC 1

RESULT 6

X12087

ID X12087 standard; DNA; 100 BP.

AC X12087;

DT 30-MAR-1999 (first entry)

DE Human biallelic polymorphic DNA fragment EST98276a.

KW Polymorphism; biallelic; human; forensic; paternity testing; disease;

KW detection; phenotypic typing; characteristic; infection; hereditary;

KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;

KW treatment; marker; ss.

OS Homo sapiens.

PN W09820165-A2.

PD 14-MAY-1998.

PF 05-NOV-1997; U20313.

PR 06-NOV-1996; US-030455.

PA (WHED) WHITEHEAD INST BIOMEDICAL RES.

PI Hudson T, Lander ES, Wang D;

DR WPI; 98-286974/25.

PT New isolated nucleic acid segments from the human genome - used for

PT determining polymorphic forms for use in e.g. forensics, paternity

PT testing or phenotypic typing for disease

PS Claim 1; Page 219; 310pp; English.

CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic

CC markers which have been isolated using the primers represented in

CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in

CC methods for determining polymorphic forms in an individual for use in

CC e.g. forensics, paternity testing or for phenotypic typing for diseases

CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,

CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial

CC hypercholesterolemia, polycystic kidney disease, hereditary

CC spherocytosis, von Willebrand's disease, tuberous sclerosis, Ehlers-Danlos

CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos

CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,

CC autoimmune diseases, inflammation, cancer, diseases of the nervous

CC system, infection by pathogenic microorganisms, and characteristics such

CC as longevity, appearance (e.g. baldness, obesity), strength, speed,

CC endurance, fertility, and susceptibility or receptivity to particular

CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid

CC segments can also be used to produce medicaments for the treatment or

CC prophylaxis of such diseases.

SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match

Best Local Similarity 0.2%; Score 63.4; DB 1; Length 100;

Matches 76; Conservative 1; Mismatches 22; Indels 0; Gaps 0;

QY 17709 GGGTCTTACTATGTTCCCGAGGCTGCTCTCAAACTCTGGCTTAAGTATCTCTCTGCC 17768

DB 1 GAGTCTTGCTTATGTTTCCCGAGGTTGCTCTTGAGCTCTGTTTCAAAACAATCTCTCTTCC 60

QY 17769 TCAGCTCCCAATGTTGGGATTACTAGTGTGAGTCAC 17807

DB 61 TAAGCTCTCTAAAGTCCCGAGGATTATAGGTGTGAGTCAC 99

RESULT 7

X12085

ID X12085 standard; DNA; 100 BP.

AC X12085;

DT 30-MAR-1999 (first entry)

DE Human biallelic polymorphic DNA fragment EST98276c.

Polymorphism; biallelic; human; forensic; paternity testing; disease;
detection; phenotypic typing; characteristic; infection; hereditary;
autoimmune disease; cancer; inflammation; drug; therapy; medication;
treatment; marker; ss.
Homo sapiens.
WO9870165-A2.
14-MAY-1998.

05-NOV-1997; U20313.
06-NOV-1996; US-030455.
(WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
WPI; 98-286974/25.

New isolated nucleic acid segments from the human genome - used for
determining polymorphic forms for use in e.g. forensics, paternity
testing or phenotypic typing for disease

Claim 1; Page 218; 310pp; English.
X10269-X12937 are human DNA fragments which contain biallelic polymorphic
markers which have been isolated using the primers represented in
X09121-X12688. The base occupying the polymorphic site is indicated by
the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
methods for determining polymorphic forms in an individual for use in
e.g. forensics, paternity testing or for phenotypic typing for diseases
such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
hypercholesterolemia, polycystic kidney disease, hereditary
spherocytosis, von Willebrand's disease, tuberculous sclerosis, Ehlers-Danlos
syndrome, osteogenesis imperfecta, acute intermittent porphyria,
autoimmune diseases, inflammation, cancer, diseases of the nervous
system, infection by pathogenic microorganisms, and characteristics such
as longevity, appearance (e.g. baldness, obesity), strength, speed,
endurance, fertility, and susceptibility or receptivity to particular
drugs or therapeutic treatments. The isolated polymorphic nucleic acid
segments can also be used to produce medicaments for the treatment or
prophylaxis of such diseases.

Sequence 100 BP; 22 A; 25 C; 22 G; 30 T;

Query Match 0.2%; Score 63.4; DB 1; Length 100;
Best Local Similarity 76.8%; Pred. No. 0.18;
Matches 76; Conservative 1; Mismatches 22; Indels 0; Gaps 0

QY 17709 GGCTCTTACTATGTGCCAGCGCTGGTCCTCAACCCFCGCGCTTAAGTGATCCTCGGCC 17768
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 GAGTCTTGTTATGTTTCCCAGGATGGCTTGAGCTGCCTGGTTTCAAACAATCTCCTCTTC 60

QY 17769 TCAGCTCCCAATCTTGGGATTACTAGTGTGAGTCAC 17807
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 TAAGCTCCYAAAGTGCCAGGATTATAGGTGTGAGTCAC 99

RESULT 8
X12086
ID X12086 standard; DNA; 100 BP.
AC X12086;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment EST98276b.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
detection; phenotypic typing; characteristic; infection; hereditary;
autoimmune disease; cancer; inflammation; drug; therapy; medication;
treatment; marker; ss.
Homo sapiens.
WO9870165-A2.
14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
(WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
WPI; 98-286974/25.

New isolated nucleic acid segments from the human genome - used for
determining polymorphic forms for use in e.g. forensics, paternity
testing or phenotypic typing for disease

Claim 1; Page 219; 310pp; English.

QY 9664 GTTAGGAGGACGC 9677
DB 14 GTTGGCAGGACG 1

RESULT 10
X12095/c

ID X12095 standard; DNA; 108 BP.
AC X12095;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment TIGR-A003M18a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN W0920165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, Ehlers-Danlos
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;

Query Match 0.2%; Score 62.2; DB 1; Length 108;
Best Local Similarity 79.4%; Pred. No. 0.27;
Matches 85; Conservative 1; Mismatches 2; Gaps 1;

QY 24415 TGTATTTCCAGCAGCTTT--GGAGGCGAGCGCGGCGAGATCCTTGTAGGTGGGAGTTCGA 24472
DB 107 TATAATCCCGAGCAGCTTTTGGGAGCGCCAGGACGAGCAGCTTGAAGTCAGGAGTTCGA 48
QY 24473 GACTAGCTTGGCCACATGATGNAACCCCATCTCTACTAAAATACA 24519
DB 47 GACCATCTCTGGCCACAYAGGAACCTCTCTCTCAAAAAAGACA 1

RESULT 11
T24892/c

ID T24892 standard; cDNA to mRNA; 100 BP.
AC T24892;
DT 05-NOV-1996 (first entry)
DE Human gene signature HUMGS06998.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.

PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PI (OKUB/) OKUBO K.
PT Matsubara K, Okubo K;
PT WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1720; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;
Query Match 0.2%; Score 60.8; DB 1; Length 100;
Best Local Similarity 74.7%; Pred. No. 0.43;
Matches 74; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 3559 TTTTCTTTTGTAGACGGAGCTCTAGCTCTGTCGCCAGGCTGAGTGCACATGCACCATC 3618
DB 100 TTTTCTTTTGTTCACAGAGATGTCACTCTGTCCACAGGCGAGTGCACCAATC 41
QY 3619 TTGGCTCACTGCAAGCTCTGCCTCCCGGTTTATGCCAT 3657
DB 40 TCAGCTNATTGCAAAATCTGCCTCCCGAGGTCAAGCGAT 2
RESULT 12
T20743/c
ID T20743 standard; cDNA to mRNA; 102 BP.
AC T20743;
DT 26-JUL-1996 (first entry)
DE Human gene signature HUMGS01961.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PI (OKUB/) OKUBO K.
PT Matsubara K, Okubo K;
PT WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 714; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-

PA (MATS/) MATSUBARA K.
PA (ORUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 1942; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 84 BP; 33 A; 17 C; 15 G; 19 T;

Query Match 0.2%; Score 56.4; DB 1; Length 84;
Best Local Similarity 80.5%; Pred. NO. 1.9;
Matches 66; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 17679 TAATTTTAAAGCTTTTCTAGAGATGGGGTCTTACTATGTTGCCAGGCTGCTC 17738

Db 82 TAATTTTAAAGCTTTTCTAGAGATGGGGTCTTACTATGTTGCCAGGCTGCTC 23

QY 17739 AAACCTCTGGGCTTAAGTGATC 17760

Db 22 GAACCTCTGGGCTCAAGGATC 1

Search completed: June 14, 2000, 20:42:29
Job time: 27860 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run On: June 14, 2000, 12:37:13 ; Search time 8513.17 Seconds
(without alignments)
13807.251 Million cell updates/sec

Title: US-08-852-495C-1_COPY_1_29000
Perfect score: 29000
Sequence: 1 CACACACACACACACACA.....CCAGGCTAGAGTCAGTGGC 29000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :	EST:*
	1: em_est1:*
	2: em_est2:*
	3: em_est3:*
	4: em_est4:*
	5: em_est5:*
	6: em_est6:*
	7: em_est7:*
	8: em_est8:*
	9: em_est9:*
	10: em_est10:*
	11: em_est11:*
	12: em_est12:*
	13: em_est13:*
	14: em_est14:*
	15: em_est15:*
	16: em_est16:*
	17: em_est17:*
	18: em_est18:*
	19: em_est19:*
	20: gb_est1:*
	21: gb_est2:*
	22: gb_est3:*
	23: gb_est4:*
	24: gb_est5:*
	25: gb_est6:*
	26: gb_est7:*
	27: gb_est8:*
	28: gb_est9:*
	29: gb_est10:*
	30: gb_est11:*
	31: gb_est12:*
	32: gb_est13:*
	33: gb_est14:*
	34: gb_est15:*
	35: gb_est16:*
	36: gb_est17:*
	37: gb_est18:*
	38: gb_est19:*
	39: gb_est20:*
	40: gb_est21:*
	41: gb_est22:*
	42: gb_est23:*
	43: gb_est24:*
	44: gb_est25:*
	45: gb_est26:*
	46: gb_est27:*
	47: gb_est28:*
	48: gb_est29:*
	49: gb_est30:*
	50: gb_est31:*
	51: gb_est32:*
	52: em_est20:*
	53: em_est21:*
	54: em_est22:*
	55: em_est23:*
	56: em_est24:*
	57: em_est25:*
	58: em_est26:*
	59: gb_est33:*
	60: gb_est34:*
	61: gb_est35:*
	62: gb_est36:*
	63: gb_est37:*
	64: gb_est38:*
	65: em_est27:*
	66: em_est28:*
	67: em_est29:*
	68: em_est30:*
	69: gb_est39:*
	70: gb_est40:*
	71: gb_est41:*
	72: gb_est42:*
	73: gb_est43:*
	74: gb_est44:*
	75: em_est31:*
	76: em_est32:*
	77: em_est33:*
	78: em_est34:*
	79: gb_est45:*
	80: gb_est46:*
	81: gb_est47:*
	82: gb_gss1:*
	83: gb_gss2:*
	84: gb_gss3:*
	85: gb_gss4:*
	86: em_gss1:*
	87: em_gss2:*
	88: em_gss3:*
	89: em_gss4:*
	90: gb_gss5:*
	91: gb_gss6:*
	92: gb_gss7:*
	93: gb_gss8:*
	94: gb_gss9:*
	95: em_gss5:*
	96: em_gss6:*
	97: em_gss7:*
	98: em_gss8:*
	99: em_gss9:*
	100: em_gss10:*
	101: em_gss11:*
	102: gb_gss10:*
	103: gb_gss11:*
	104: em_gss12:*
	105: gb_gss12:*
	106: gb_gss13:*
	107: gb_gss14:*
	108: gb_gss15:*
	109: gb_gss16:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query
--------	-------

clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 727 Std Error: 0.00

Seq primer: M13RP1

High quality sequence stop: 68.

FEATURES

Location/Qualifiers

1. .95

/organism="Homo sapiens"

/db_xref="taxon:9606"

/db_xref="taxon:9606"

/clone="IMAGE:127301"

/clone_lib="Soares fetal liver spleen INFUS"

/sex="male"

/dev_stage="20 week-post conception fetus"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)

with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;

1st strand cDNA was primed with a Pac I - oligo(dT) primer

15' AACTGGAAGAAATAAAGATCTTTTTTTTTTTTTTTT 3'],

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Pac I and cloned into the Pac I

and Eco RI sites of the modified pT7T3 vector. Library

went through one round of normalization. Library

constructed by Bento Soares and M.Fatima Bonaldo."

31 a 23 c 26 g 12 t 3 others

BASE COUNT 31 a 23 c 26 g 12 t 3 others

ORIGIN

Query Match 0.3%; Score 92; DB 21; Length 95;

Best Local Similarity 96.8%; Pred. No. 0.066;

Matches 92; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 10316 TGTGTTAAGTCTCAGCAGCAGCAAGTCTGCTCCGCCGCGGAGGAGGCTCCCAAG 10375

Db 1 TGTGTTAAGTCTCAGCAGCAGCAAGTCTGCTCCGCCGCGGAGGAGGCTCCCAAG 60

QY 10376 GCAGTGACCAAGCGCAGAGAAAGATGGCAAGAA 10410

Db 61 GCAGTGACCAAGCGCAGAGAAAGATGGCAAGAA 95

RESULT 3

AI832832

LOCUS 105 bp mRNA EST 13-JUL-1999

DEFINITION at72g09.x1 Barstead colon HPLR87 Homo sapiens cDNA clone

IMAGE:2377600 3' similar to contains Alu repetitive

element;contains element MER22 repetitive element ;, mRNA sequence.

ACCESSION AI832832

VERSION AI832832.1 GI:5454812

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 105)

AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,

Krisman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,

Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,

Theising,B., White,Y., Wyllie,T., Waterston,R. and Wilson,R.

WashU-NCI human EST Project

Unpublished (1997)

On Dec 20, 1995 this sequence version replaced gi:1133644.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -400P from Glibco.

High quality sequence stop: 68.

Location/Qualifiers

1. .105

FEATURES

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2377600"

/clone_lib="Barstead colon HPLR87"

/sex="male"

/dev_stage="adult, age 25"

/lab_host="DH10B (phage resistant)"

/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a

modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer [5'

TGTTACGATCTGAAGTGGAGCGCCGCTTTTTTTTTTTTTTTT

3']; double-stranded cDNA was ligated to Eco RI adaptors

[5' AATTCAGTAGTAAT 3' and 5' ATTACTAGTG 3'], digested

with Not I and cloned into the Not I and Eco RI sites of

the modified pT7T3 vector. Library constructed by Bob

Barstead."

17 a 35 c 27 g 26 t

BASE COUNT 17 a 35 c 27 g 26 t

ORIGIN

Query Match 0.3%; Score 92.2; DB 61; Length 105;

Best Local Similarity 92.4%; Pred. No. 0.06;

Matches 97; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 12176 GAGACCAAGTTTCTCTGTTTCCAGGCTGGAGTGCATGCGCGATCTTGGCTCAC 12235

Db 1 GAGACAGAGTTTGGCTCTGTTGTCGCCAGGCTGGAGTGCATGCGCGATCTTGGCTCAC 60

QY 12236 GCACCTCCGCTCCGCGGTTCAAGCAATTCCTCCTCGCTCAGCCT 12280

Db 61 GCACCTCCGCTCCGCGGTTCAAGCAATTCCTCCTCGCTCAGCCT 105

RESULT 4

AA807640

LOCUS 103 bp mRNA EST 05-MAR-1998

DEFINITION nx08b05.s1 NCI_CGAP_GC3 Homo sapiens cDNA clone IMAGE:1255473 3'

similar to contains Alu repetitive element;; mRNA sequence.

ACCESSION AA807640

VERSION AA807640.1 GI:2877108

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 103)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

On Jan 19, 1998 this sequence version replaced gi:2151346.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert_Strausberg@nih.gov

Tissue procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 774 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 87.

Location/Qualifiers

1. .103

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1255473"

/clone_lib="NCI_CGAP_GC3"

/tissue_type="pooled germ cell tumors"

FEATURES

source

```

/lab_host="DH10B"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is not normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      19 a      27 c      30 g      27 t
ORIGIN

Query Match      0.3%; Score 87.6; DB 38; Length 103;
Best Local Similarity 91.2%; Pred. No. 0.21;
Matches 93; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3734 AGTAGAGATGGGGTTTACCGTGTAGCAGATGGTCTCGATCTCTTGTGATC 3793
      |||||
Db 2 AGTAGAGATGGGGTTTACCGTGTAGCAGATGGTCTCGATCTCTTGTGATC 61
      |||||

QY 3794 CGCCTGGCTTGCCTCCCAAGTCTGGGATACACGTGTGA 3835
      |||||
Db 62 CGCTCACCTCGCCCTCCCAAGTCTGGGATACAGGTGTGA 103
      |||||

RESULT 5
T77382/c
LOCUS T77382 103 bp mRNA EST 15-MAR-1995
DEFINITION yD72h12.r1 Soares fetal liver spleen lNFLS Homo sapiens cDNA clone
IMAGE:113831 5' similar to contains Alu repetitive element; mRNA
sequence.
ACCESSION T77382
VERSION T77382.1 GI:694585
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 103)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Luktan, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Other_ESTS: yD72h12.s1
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 943
Source: IMAGE Consortium, LLNL This clone is available royalty-free
through LLNL; contact the IMAGE Consortium (info@image.llnl.gov)
for further information. Putative full length read
Insert Length: 943 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 109.
Location/Qualifiers
1. 103
/organism="Homo sapiens"
/db_xref="db:469448"
/db_xref="taxon:9606"
/clone="IMAGE:113831"
/clone_lib="Soares fetal liver spleen lNFLS"
/sex="male"
/dev_stage="20 week post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;

```

```

1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' ACTGGAAGAAATTAATAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      24 a      20 c      37 g      22 t
ORIGIN

Query Match      0.3%; Score 86.6; DB 21; Length 103;
Best Local Similarity 95.7%; Pred. No. 0.28;
Matches 89; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6077 TGAGAGTCTCACTCACTGCAACCTCCCTCTCTATATTCAGTGATCTTGGCTCA 6136
      |||||
Db 103 TGAGAGTCTCACTCACTGCAACCTCCCTCTCTATATTCAGTGATCTTGGCTCA 44
      |||||

QY 6137 GCCTCCCGAGTAGCTGGGACTACAGCGCTGCAC 6169
      |||||
Db 43 GCCTCCCGAGTAGCTGGGACTACAGCGCTGCAC 11
      |||||

RESULT 6
AA158786/c
LOCUS AA158786 106 bp mRNA EST 09-MAR-1998
DEFINITION z663cll.r1 Stragatene pancreas (#937208) Homo sapiens cDNA clone
IMAGE:591572 5' similar to contains Alu repetitive element; contains
element PTR7 repetitive element; mRNA sequence.
ACCESSION AA158786
VERSION AA158786.1 GI:1733588
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 106)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1406940.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Putative full length read
The vector to vector length is 119
Insert Length: 926 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham.
Location/Qualifiers
1. 106
/organism="Homo sapiens"
/db_xref="db:462258"
/db_xref="taxon:9606"
/clone="IMAGE:591572"
/clone_lib="Stragatene pancreas (#937208)"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: pancreas; Vector: Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Pancreatic adenocarcinoma cell line. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor
sequence: 5' GAATTCGACGAG 3' -3' adaptor sequence: 5'
CTCGAGTCTTTTTTTTTTTTTTTT 3"
BASE COUNT      27 a      28 c      37 g      14 t
ORIGIN

```

```

Query Match      0.3%; Score 84.8; DB 29; Length 106;
Best Local Similarity 88.5%; Pred. No. 0.45; Mismatches 0; Gaps 0;
Matches 92; Conservative 0;

QY 12190 CTCCTGTTCCAGGTCGAGTGCAATGGCGGATCTTGGCTCACAGCAACCTCCGCCTC 12249
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 104 CTCCTGTTCCAGGTCGAGTGCAATGGCGGATCTTGGCTCACAGCAACCTCCGCCTC 45
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 12250 CCGGGTTCAGGCATCTCTCGCTCAGCTCCGGAGTAGCTGG 12293
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44 CCAAGTTAAGCATCTCTGCTGCCCGCCTCTCTGAGTGGCTGG 1

RESULT 7
AQ029690/c
LOCUS
DEFINITION
  RPc11-41f18-TV RPc1-11 Homo sapiens genomic clone RPc1-11-41f18,
  genomic survey sequence.
ACCESSION
  AQ029690
VERSION
  AQ029690.1 GI:3274821
KEYWORDS
  GSS.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 109)
AUTHORS
  Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
  Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
  Venter,J.C.
  Use of BAC End Sequences for Sequence-Ready Map Building (1998)
  Unpublished (1998)
JOURNAL
  Department of Mark Adams
COMMENT
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: mdamads@tigr.org
  Clones are derived from the human BAC library RPc1-11. For BAC
  library availability, please contact Pieter de Jong
  (pieter@dejong.med.buffalo.edu). Clones may be purchased from
  BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
  Research Genetics (info@resgen.com). BAC end search page:
  http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
  Seq primer: T7
  Class: BAC ends.
  Location/Qualifiers
    1..109
    /organism="Homo sapiens"
    /db_xref="GDB:7515497"
    /db_xref="taxon:9606"
    /clone="RPc1-11-41f18"
    /clone_lib="RPc1-11"
    /sex="Male"
    /cell_type="Lymphocytes"
    /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
    RPc11 Human Male BAC Library"
  BASE COUNT      21 a 26 c 24 g 38 t
  ORIGIN

Query Match      0.3%; Score 85; DB 94; Length 109;
Best Local Similarity 86.2%; Pred. No. 0.42;
Matches 94; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 4703 CATATCACCTGAGGTGAGAGTTGAGACGAGCTGCCCAACATGTTGAACCCCTGTC 4762
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 109 CAGATCACCTGAGGTGAGAGTTGAGACGAGCTGCCCAACATGTTGAACCCCTGTC 50
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 4763 TACTATAAATATAAAATTAAGCTGGGTGGTGGTGCATGCGCTGTAGTC 4811
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 49 TACTAAACTACAAAATTTAGCCGGCATGAAGGAGCATGACTGTAAATC 1

RESULT 8
AA244245
LOCUS
DEFINITION
  AA244245 110 bp mRNA EST 20-AUG-1997
  nc07a04.s1 NCI_CGAP_Prl Homo sapiens cDNA clone IMAGE:1007406
  similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION
  AA244245
VERSION
  AA244245.1 GI:1875104
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 110)
AUTHORS
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL
  Unpublished (1997)
COMMENT
  On Jan 24, 1995 this sequence version replaced gi:634306.
  Contact: Robert Strausberg, Ph.D.
  Tel: (301) 496-1550
  Email: Robert_Strausberg@nih.gov
  Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuauqui,
  M.D., Michael Emmert-Buck, M.D., Ph.D.
  CDNA Library Preparation: David B. Krizman, Ph.D.
  CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
  DNA sequencing by: Washington University Genome Sequencing Center
  clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -41ml3 fwd. ET from Amersham
High quality sequence stop: 90.
  Location/Qualifiers
    1..110
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:1007406"
    /clone_lib="NCI_CGAP_Prl"
    /sex="Male"
    /dev_stage="45 years old"
    /lab_host="DH10B"
    /note="Vector: PAMP10; Site_1: NotI; Site_2: EcoRI; 1st
    strand cDNA was primed with oligo(GT)17 on 50 ng of
    DNase-treated, total cellular RNA obtained from
    5,000-10,000 microdissected, histologically normal
    prostate epithelial cells. Double-stranded cDNA was
    ligated to EcoRI adaptors, 5 cycles of PCR applied to the
    cDNA with an adaptor-specific primer, and the resulting
    PCR product subcloned into pAMP10 by the UDG-cloning
    method (Life Technologies). Average insert size is 600
    bp. NOTE: Not directionally cloned. This library was
    constructed by David Krizman."
  BASE COUNT      17 a 26 c 28 g 38 t
  ORIGIN

Query Match      0.3%; Score 85; DB 30; Length 110;
Best Local Similarity 85.5%; Pred. No. 0.42;
Matches 94; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 3560 TTTTCTTTTGGACGAGCTAGCTCTGTGCGCCAGCTGGAGTCAGTGGCACCATCT 3619
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TTTTCTTTTGGAGTGGAGTCTGTGTTGCCAGCTGGAGTCAGTGGCAGANTCT 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 3620 TGGCTCACATGCAGCTCTGCTCCCGGTTATGCCATCTCATGTCCTCA 3669
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TGGCTCACATGCAGCTCTGCTCCCGGTTCAAGAGATTCTTCTGCTCTCA 110
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9

```

```

AA897366
LOCUS      AA897366      110 bp      mRNA      EST      04-JAN-1999
DEFINITION am06h02.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:1466067 3' similar to contains Alu repetitive element;; mRNA
sequence.

ACCESSION AA897366.1 GI:3033986
VERSION    AA897366
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 110)
AUTHORS    Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and
            Venter, J.C.
TITLE      Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
JOURNAL    Unpublished (1997)
COMMENT    Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbeetigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pietred@jmg.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: T7
            Class: BAC ends.
            Location/Qualifiers
                source
                    1..110
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:1466067"
                    /clone_11b="Soares_NFL_T_GBC_S1"
                    /lab_host="DH10B"
                    /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                    a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
                    Equal amounts of plasmid DNA from three normalized
                    libraries (fetal lung NBHL19W, testis NHT, and B-cell
                    NCI CGAP-GCBI) were mixed, and ss circles were made in
                    vitro. Following HAP purification, this DNA was used as
                    tracer in a subtractive hybridization reaction. The driver
                    was PCR-amplified cDNAs from pools of 5,000 clones made
                    from the same 3 libraries. The pools consisted of
                    I.M.A.G.E. clones 297480-302087, 682632-687239,
                    726408-728711, and 729096-731399. Subtraction by Bento
                    Soares and M. Fatima Bonaldo."
BASE COUNT 22 a 27 c 29 g 32 t
ORIGIN
Query Match 0.3%; Score 85; DB 39; Length 110;
Best Local Similarity 86.2%; Pred. No. 0.42;
Matches 94; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 3726 TTTATTATTAGTAGAGATGGGGTTTCCCGTGTACCCAGATGGTCTCGATCTCTTGACC 3785
+ + + + +
Db 2 TATTTTATTAGTAGATGGGGTTTCCCGTGTACCCAGATGGTCTCAATCTCTGGACC 61
+ + + + +

QY 3786 TTCTGTCGGCTGCTTGGCTTCCCAAGTCTGGGATTACACGTGTG 3834
+ + + + +
Db 62 TCATGATCGGCCACCTCGGCTCCCAAGTCTGGGATTATAGGCGTG 110
+ + + + +

RESULT 10
LOCUS      AQ535244/c
DEFINITION RPCI-11-317H22.IV RPCI-11 Homo sapiens genomic clone
IMAGE:1140858 5' similar to contains Alu repetitive element;; mRNA
sequence.
ACCESSION AQ535244
VERSION    AQ535244.1 GI:4846934
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 106)
AUTHORS    Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisels, G., Jost, S.,
            Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
            Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
            Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
            WashU-NCI human EST Project
            Unpublished (1997)
            On Sep 12, 1996 this sequence version replaced gi:1397630.
            Contact: Wilson RK
            Washington University School of Medicine

Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 103)
Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and
Venter, J.C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
Unpublished (1997)
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pietred@jmg.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
    source
        1..103
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:1466067"
        /clone_11b="Soares_NFL_T_GBC_S1"
        /lab_host="DH10B"
        /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
        a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
        Equal amounts of plasmid DNA from three normalized
        libraries (fetal lung NBHL19W, testis NHT, and B-cell
        NCI CGAP-GCBI) were mixed, and ss circles were made in
        vitro. Following HAP purification, this DNA was used as
        tracer in a subtractive hybridization reaction. The driver
        was PCR-amplified cDNAs from pools of 5,000 clones made
        from the same 3 libraries. The pools consisted of
        I.M.A.G.E. clones 297480-302087, 682632-687239,
        726408-728711, and 729096-731399. Subtraction by Bento
        Soares and M. Fatima Bonaldo."
BASE COUNT 22 a 27 c 29 g 32 t
ORIGIN
Query Match 0.3%; Score 84.4; DB 108; Length 103;
Best Local Similarity 89.2%; Pred. No. 0.5;
Matches 91; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 12327 TTTTGTATTATTAGTAGAGACGAGGTTTCTCCATGTCGTCAGGTCGTCTCGAACCTCG 12386
+ + + + +
Db 102 TTTTGTATTATTAGTAGAGACGAGGTTTCTCCATGTCGTCAGGTCGTCTCGAACCTCT 43
+ + + + +

QY 12387 GACATCAGGTGATGTCGCCGCTTGGCTCCCAAGTCTCGT 12428
+ + + + +
Db 42 GACCTCAAGTGATCTGCCGCTTGGCTCCCAAGTCTCGT 1
+ + + + +

RESULT 11
LOCUS      AA703692
DEFINITION ag81a10.r1 Stratagene hMT neuron (#937233) Homo sapiens cDNA clone
IMAGE:1140858 5' similar to contains Alu repetitive element;; mRNA
sequence.
ACCESSION AA703692
VERSION    AA703692.1 GI:2713610
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 106)
AUTHORS    Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisels, G., Jost, S.,
            Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
            Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
            Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
            WashU-NCI human EST Project
            Unpublished (1997)
            On Sep 12, 1996 this sequence version replaced gi:1397630.
            Contact: Wilson RK
            Washington University School of Medicine

```


Db 61 CATGAATTCCTTCGTCACGACATGTCGAGCGCATC 97

Search completed: June 14, 2000, 15:22:36
Job time: 9923 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: June 14, 2000, 12:54:30 ; Search time 372.68 seconds
(without alignments)
10114.754 Million cell updates/sec

Title: US-08-852-495C-1_COPY_1_29000

Perfect score: 29000

Sequence: 1 CACACACACACACACACACA.....CCAGGCTAGAGTCAGTGGC 29000

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10

Maximum DB seq length: 110

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/lna/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/lna/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/lna/5C_COMB.seq:*
- 4: /cgn2_6/ptodata/1/lna/5D_COMB.seq:*
- 5: /cgn2_6/ptodata/1/lna/6_COMB.seq:*
- 6: /cgn2_6/ptodata/1/lna/PCITUS_COMB.seq:*
- 7: /cgn2_6/ptodata/1/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	71.4	0.2	105	4	US-08-481-658B-65
2	71.4	0.2	105	4	US-08-477-504A-65
3	71.4	0.2	105	4	US-08-486-756A-65
4	71.4	0.2	105	4	US-08-485-862B-65
5	71.4	0.2	105	5	US-08-787-739-65
6	69.6	0.2	105	4	US-08-481-658B-65
7	69.6	0.2	105	4	US-08-477-504A-65
8	69.6	0.2	105	4	US-08-486-756A-65
9	69.6	0.2	105	4	US-08-485-862B-65
10	69.6	0.2	105	5	US-08-787-739-65
11	58.6	0.2	78	3	US-08-454-557C-70
12	58.6	0.2	78	4	US-08-340-426D-70
13	58.6	0.2	78	4	US-08-450-673C-70
14	58.6	0.2	78	6	PCIT-US95-17111A-70
15	57.2	0.2	78	3	US-08-454-557C-70
16	57.2	0.2	78	4	US-08-340-426D-70
17	57.2	0.2	78	4	US-08-450-673C-70
18	57.2	0.2	78	6	PCIT-US95-17111A-70
19	55.8	0.2	84	3	US-08-454-557C-91
20	55.8	0.2	84	4	US-08-340-426D-91
21	55.8	0.2	84	4	US-08-450-673C-91
22	55.8	0.2	84	6	PCIT-US95-17111A-91
23	54.8	0.2	76	3	US-08-454-557C-69
24	54.8	0.2	76	4	US-08-340-426D-69
25	54.8	0.2	76	4	US-08-450-673C-69
26	54.8	0.2	76	6	PCIT-US95-17111A-69
27	55	0.2	84	3	US-08-454-557C-91

Sequence 91, Appl
Sequence 91, Appl
Sequence 91, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 92, Appl
Sequence 92, Appl
Sequence 92, Appl
Sequence 36, Appl
Sequence 36, Appl
Sequence 57, Appl
Sequence 57, Appl
Sequence 57, Appl
Sequence 60, Appl
Sequence 60, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-65
; Sequence 65, Application US/08481658B
; Patent No 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.2%; Score 71.4; DB 4; Length 105;
Best Local Similarity 80.0%; Pred. No. 5.2e-07;

Matches 84; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 3720 TTTTATTTTATTTAGATAGAGTGGGTTTCACCGTGTAGCCAGAAAGTCTCGATCTC 3779

Db 1 TTTTATACATCTTTAGTAGACAGAGGTTTACCATATATGGCCAGGCTGCTCTCAAACTC 60

Qy 3780 TTGACCTTCTGATCGCGCTGCTTGGCTTCCCAAAGTCTGGGAT 3824

Db 61 CTGACCTTGTGATCCACCAGCGCTCGGCCCTCCCAAAGTCTGGGAT 105

RESULT 2

US-08-477-504A-65

; Sequence 65, Application US/08477504A

; Patent No. 5972353

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,504A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3D

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-477-504A-65

Query Match 0.2%; Score 71.4; DB 4; Length 105;

Best Local Similarity 80.0%; Pred. No. 5.2e-07;

Matches 84; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 3720 TTTTATTTTATTTAGATAGAGTGGGTTTCACCGTGTAGCCAGAAAGTCTCGATCTC 3779

Db 1 TTTTATACATCTTTAGTAGACAGAGGTTTACCATATATGGCCAGGCTGCTCTCAAACTC 60

Qy 3780 TTGACCTTCTGATCGCGCTGCTTGGCTTCCCAAAGTCTGGGAT 3824

Db 61 CTGACCTTGTGATCCACCAGCGCTCGGCCCTCCCAAAGTCTGGGAT 105

RESULT 3

US-08-486-756A-65

; Sequence 65, Application US/08486756A

; Patent No. 5981711

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,756A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-486-756A-65

Query Match 0.2%; Score 71.4; DB 4; Length 105;

Best Local Similarity 80.0%; Pred. No. 5.2e-07;

Matches 84; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 3720 TTTTATTTTATTTAGATAGAGTGGGTTTCACCGTGTAGCCAGAAAGTCTCGATCTC 3779

Db 1 TTTTATACATCTTTAGTAGACAGAGGTTTACCATATATGGCCAGGCTGCTCTCAAACTC 60

Qy 3780 TTGACCTTCTGATCGCGCTGCTTGGCTTCCCAAAGTCTGGGAT 3824

Db 61 CTGACCTTGTGATCCACCAGCGCTCGGCCCTCCCAAAGTCTGGGAT 105

RESULT 4

US-08-485-862B-65

; Sequence 65, Application US/08485862B

; Patent No. 5989838

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,862B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-485-862B-65

Query Match 0.2%; Score 71.4; DB 4; Length 105;
Best Local Similarity 80.0%; Pred. No. 5.2e-07;
Matches 84; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3720 TTTTATTTTATTTAGTAGAGATGGGTTTACCGTGTAGCCAGAAATGCTCTCGATCTC 3779
Db 1 TTTTATTTTATTTAGTAGAGATGGGTTTACCGTGTAGCCAGAAATGCTCTCTCAAACTC 60

QY 3780 TTGACCTTCTGATCCCGCTGGCTTCCCAAAAGTCTGGGAT 3824
Db 61 CTGACCTTGTGATCCAGCAGCTCGGCTCCCAAAAGTCTGGGAT 105

RESULT 5
US-08-787-739-65
; Sequence 65, Application US/08/787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/787,739
FILING DATE: 24-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/486,756
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/481,658
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,862
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,863
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,077
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-981-0332
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-787-739-65

Query Match 0.2%; Score 71.4; DB 5; Length 105;
Best Local Similarity 80.0%; Pred. No. 5.2e-07;
Matches 84; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3720 TTTTATTTTATTTAGTAGAGATGGGTTTACCGTGTAGCCAGAAATGCTCTCGATCTC 3779
Db 1 TTTTATTTTATTTAGTAGAGATGGGTTTACCGTGTAGCCAGAAATGCTCTCTCAAACTC 60

QY 3780 TTGACCTTCTGATCCCGCTGGCTTCCCAAAAGTCTGGGAT 3824
Db 61 CTGACCTTGTGATCCAGCAGCTCGGCTCCCAAAAGTCTGGGAT 105

RESULT 6
US-08-481-658B-65/c
; Sequence 65, Application US/08/481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481.658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.2%; Score 69.6; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. No. 1.3e-06;
Matches 89; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

Qy 1395 ATCTCAGACACTTTGGGAGGCTGAGG-GCACAGATCAGAGGTGGGAGTTTGAGACCAGC 1453
||| |||||||||||||||| |||| |||| |||| |||||||||||| ||||
Db 105 ATCCAGCAGCACTTTGGGAGCGGAGCTGGTGATCACAAGGTCAGGAGTTTGAGAGCAGC 46
||| |||||||||||||||| |||| |||| |||| |||||||||||| ||||

Qy 1454 CTGGCCAATATGGCAACCCCTGCTCTACTAAATAACAAAA 1497
||||| |||||| |||||| |||||| |||||| |||||| || |||||
Db 45 CTGGCCAATATGGTGAACCCCTGCTCTACTAAAGATGTAATA 2

RESULT 7
US-08-477-504A-65/c
Sequence 65, Application US/08477504A
Patent No. 5972353
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477.504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-477-504A-65

Query Match 0.2%; Score 69.6; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. No. 1.3e-06;
Matches 89; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

Qy 1395 ATCTCAGACACTTTGGGAGGCTGAGG-GCACAGATCAGAGGTGGGAGTTTGAGACCAGC 1453
||| |||||||||||||||| |||| |||| |||| |||||||||||| ||||
Db 105 ATCCAGCAGCACTTTGGGAGCGGAGCTGGTGATCACAAGGTCAGGAGTTTGAGAGCAGC 46
||| |||||||||||||||| |||| |||| |||| |||||||||||| ||||

Qy 1454 CTGGCCAATATGGCAACCCCTGCTCTACTAAATAACAAAA 1497
||||| |||||| |||||| |||||| |||||| |||||| || |||||
Db 45 CTGGCCAATATGGTGAACCCCTGCTCTACTAAAGATGTAATA 2

RESULT 8
US-08-486-756A-65/c
Sequence 65, Application US/08486756A
Patent No. 5981711
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486.756A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-486-756A-65

Query Match 0.2%; Score 69.6; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. No. 1.3e-06;
Matches 89; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 1395 ATCTCAGCACTTTGGGAGGCTGAGG-GCACAGATCACGAGTCGGGAGTTTGAGACCAGC 1453
DB 105 ATCCCGAGCACTTTGGGAGCGCCGAGGCTGGTGGATCAAGGTCAGGAGTTTGAGACCAGC 46

QY 1454 CTGGCCAATATGCGGAACCCCTCTCTCTACTAAATACAAAA 1497
DB 45 CTGGCCAATATGCTGAACCCCTCTCTCTACTAAAGATGTAAAAA 2

RESULT 9

US-08-485-862B-65/C
Sequence 65, Application US/08485862B
Patent No. 5989838

GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:

ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA

ZIP: 94920

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,862B

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/477,504

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/260,190

FILING DATE: 15-JUN-1994

ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.

REGISTRATION NUMBER: 30,863

REFERENCE/DOCKET NUMBER: D-0021.3D

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-435-2034

TELEFAX: 415-435-0727

INFORMATION FOR SEQ ID NO: 65:

SEQUENCE CHARACTERISTICS:

LENGTH: 105 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-485-862B-65

Query Match 0.2%; Score 69.6; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. No. 1.3e-06;
Matches 89; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 1395 ATCTCAGCACTTTGGGAGGCTGAGG-GCACAGATCACGAGTCGGGAGTTTGAGACCAGC 1453
DB 105 ATCCCGAGCACTTTGGGAGCGCCGAGGCTGGTGGATCAAGGTCAGGAGTTTGAGACCAGC 46
QY 1454 CTGGCCAATATGCGGAACCCCTCTCTCTACTAAATACAAAA 1497
DB 45 CTGGCCAATATGCTGAACCCCTCTCTCTACTAAAGATGTAAAAA 2

RESULT 10

US-08-787-739-65/C

Sequence 65, Application US/08787739

Patent No. 6027387

GENERAL INFORMATION:

APPLICANT: Zavada, Jan

APPLICANT: Pastorekova, Silvia

APPLICANT: Pastorek, Jaromir

TITLE OF INVENTION: MN Gene and Protein

NUMBER OF SEQUENCES: 96

CORRESPONDENCE ADDRESS:

ADDRESSEE: Leona L. Lauder

STREET: 369 Pine Street, Suite 610

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/787,739

FILING DATE: 24-JAN-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/485,049

FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/486,756

FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/477,504

FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/481,658

FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/485,862

FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/485,863

FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/487,077

FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.

REGISTRATION NUMBER: 30,863

REFERENCE/DOCKET NUMBER: D-0021.4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-981-2034

TELEFAX: 415-981-0332

INFORMATION FOR SEQ ID NO: 65:

SEQUENCE CHARACTERISTICS:

LENGTH: 105 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-787-739-65

```
Query Match          0.2%; Score 69.6; DB 5; Length 105;
Best Local Similarity 85.6%; Pred. No. 1.3e-06;
Matches 89; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

Qy 1395 ATCTCAGCAGCTTTGGGAGGCTGAGG-GCACAGATCACAGGTGGGAGTTTGACAGCAGC 1453
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 105 ATCCAGCAGCAGCTTTGGGAGGCGGAGGCTGGTGGATCACAGGTTCAGGAGTTTGAGAGCAGC 46

Qy 1454 CTGGCCAATATGCGAAACCCCTGCTCTACTAAATAACAAAAA 1497
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 45 CTGGCCAATATGCTGAAACCCCTGCTCTACTAAAGATGTAATAA 2

RESULT 11
US-08-454-557C-70
; Sequence 70, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 30-MAY-1995
; APPLICATION NUMBER: US/08/454,557C
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2540
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-454-557C-70

Query Match          0.2%; Score 58.6; DB 3; Length 78;
Best Local Similarity 87.7%; Pred. No. 0.00029;
Matches 64; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 2840 GCCCAGCTAATTTTGTATTTTAGTAGAGATGGGGTTTCTACTATGTGGCCAGGCTAGT 2899
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6 GCCCAGCTAATTTTGTATTTTAGTAGAGATGGGGTTTCTCTCCATGTTTCATCAGGCTGTT 65

Qy 2900 TTGGAACCTCTCA 2912
    | ||||| |||||
Db 66 GTCGAACCTCTCA 78

RESULT 13
US-08-450-673C-70
; Sequence 70, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 14-NOV-1994
; APPLICATION NUMBER: US/08/340,426D
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-340-426D-70

Query Match          0.2%; Score 58.6; DB 4; Length 78;
Best Local Similarity 87.7%; Pred. No. 0.00029;
Matches 64; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 2840 GCCCAGCTAATTTTGTATTTTAGTAGAGATGGGGTTTCTACTATGTGGCCAGGCTAGT 2899
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6 GCCCAGCTAATTTTGTATTTTAGTAGAGATGGGGTTTCTCCATGTTTCATCAGGCTGTT 65

Qy 2900 TTGGAACCTCTCA 2912
    | ||||| |||||
Db 66 GTCGAACCTCTCA 78

RESULT 13
US-08-450-673C-70
; Sequence 70, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```

; TOPOLOGY: both
PCT-US95-17111A-70

Query Match          0.2%   Score 58.6;   DB 6;   Length 78;
Best Local Similarity 87.7%;   Pred No. 0.00029;
Matches 64;   Conservative 0;   Mismatches 9;   Indels 0;   Gaps 0;

QY  2840  GCCCAGCTAAATTTTGTGATTTTTAGTAGAGATGGGTTTCACTATGTTGGCCAGGCTAGT 2899
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    6    GCCCAGCTAAATTTGTGATTTTTAGTAGAGATGGGTTTCTCCATGTTTCATCAGGCTGCT 65

QY  2900  TTGGAACCTCCTGA 2912
      | ||||| |||||
Db    66  GTCGAACCTCCTGA 78

RESULT 15
US-08-454-557C-70/G
; Sequence 70, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: Of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington

```

```

? COUNTRY: U.S.A.
? ZIP: 20005-3934
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/454.557C
? FILING DATE: 30-MAY-1995
? CLASSIFICATION: 51A
? ATTORNEY/AGENT INFORMATION:
? NAME: Ludwig, Steven R.
? REGISTRATION NUMBER: 36,203
? REFERENCE/DOCKET NUMBER: 0609.3840003
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 371-2600
? TELEFAX: (202) 371-2540
? INFORMATION FOR SEQ ID NO: 70:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 78 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: both
? TOPOLOGY: both
?
? US-08-454-557C-70

```

```

Query Match          0.2%; Score 57.2; DB 3; Length 78;
Best Local Similarity 83.3%; Pred. No. 0.00059;
Matches 65; Conservative 0; Mismatches 13; Indels 0; Gaps 0

QY 1435 TCGGAGTTTCAGACCCAGCTGGCCAATATGCGGAACCCTGTCTCTACTAAATAACA 1494
Db      TCAGGATTTCAGCACACGCCAGCTGATGAACATGGAGAACCCTCATCTCTACTAAATAACA 19

QY 1495 AAATTAGCTGGGCATGGT 1512
Db      | ||||| ||||| |||||
        18 ATATTAGCTGGCGTGTT 1
US-08-454-557C-70

```

Search completed: June 14, 2000, 20:28:41
Job time: 27251 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2000, 20:22:29 ; Search time 17971.3 Seconds
(without alignments)
-1569.830 Million cell updates/sec

Title: US-08-852-495C-1_COPY_28000_57000
Perfect score: 29001
Sequence: 1 ATATCAACAAAAACACACAT.....TTAGCAGCACACAGGTAGGCT 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba1.*
2: gb_ba2.*
3: gb_on.*
4: gb_ov.*
5: gb_pat.*
6: gb_ph.*
7: gb_pl1.*
8: gb_pl2.*
9: gb_pr1.*
10: gb_pr2.*
11: gb_pr3.*
12: gb_ro.*
13: gb_sts.*
14: gb_sy.*
15: gb_un.*
16: gb_vi.*
17: em_fun.*
18: em_hum1.*
19: em_hum2.*
20: em_in.*
21: em_on.*
22: em_or.*
23: em_ov.*
24: em_pat.*
25: em_ph.*
26: em_pl.*
27: em_ro.*
28: em_sts.*
29: em_sy.*
30: em_un.*
31: em_vi.*
32: gb_htg1.*
33: gb_htg2.*
34: gb_in1.*
35: gb_in2.*
36: em_ba1.*
37: em_ba2.*
38: em_hum3.*
39: em_hum4.*
40: gb_pr4.*
41: gb_htg3.*
42: gb_htg4.*
43: gb_htg5.*
44: gb_htg6.*

45: gb_htg7.*
46: em_htg1.*
47: em_htg2.*
48: em_htg3.*
49: em_hum5.*
50: gb_pl3.*
51: gb_pr5.*
52: gb_htg8.*
53: gb_htg9.*
54: gb_htg10.*
55: gb_htg11.*
56: gb_htg12.*
57: gb_htg13.*
58: gb_htg14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	94.4	0.3	96	13	G31304	G31304 sy899gl-19
2	91.6	0.3	108	10	HSLDLRN2	X05250 Human LDL-r
3	90.4	0.3	108	10	HSLDLRN2	X05250 Human LDL-r
4	88	0.3	108	11	HSU67803	U67803 Human small
5	87	0.3	107	9	HUMALCE162	M87924 Human carci
6	87	0.3	108	11	HSU67804	U67804 Human small
7	85.4	0.3	103	9	HUMALCE221	M87896 Human carci
8	84	0.3	103	9	HUMALCE221	M87896 Human carci
9	84.2	0.3	108	10	HSLDLRD1	X05249 Human LDL-r
10	84.2	0.3	108	10	HSLDLRD2	X05251 Human LDL-r
11	83	0.3	108	10	HSLDLRD1	X05249 Human LDL-r
12	83	0.3	108	10	HSLDLRD2	X05251 Human LDL-r
13	81.4	0.3	108	10	HSLDLI12	X05248 Human LDL-r
14	79	0.3	108	11	HSU67808	U67808 Human small
15	78.8	0.3	110	9	HUMALCE43	M87900 Human carci
16	78.2	0.3	110	11	HSU67807	U67807 Human small
17	77.4	0.3	90	9	HUMLDLRFL	X03555 Human low d
18	77.6	0.3	103	13	HS8IC8R	X57789 Human sequ
19	77.6	0.3	104	9	HUMALCE272	M87899 Human carci
20	77.4	0.3	107	9	HUMALCE162	M87924 Human carci
21	77.2	0.3	108	10	HSLDLI12	X05248 Human LDL-r
22	76.2	0.3	108	11	HSU67803	U67803 Human small
23	75.4	0.3	97	9	HUMLDLRDJ	M14179 Human fami
24	75.6	0.3	103	13	HS8IC8R	X57789 Human sequ
25	75	0.3	110	11	HSU67807	U67807 Human small
26	73.8	0.3	97	9	HUMLDLRA1	M14178 Human low d
27	73.6	0.3	97	9	HUMLDLRA2	M14180 Human low d
28	73	0.3	107	11	HSU67806	U67806 Human small
29	73.2	0.3	110	9	HUMALCE43	M87900 Human carci
30	72.6	0.3	100	9	HUMGALNSA	D45223 Human GALNS
31	72.6	0.3	108	9	HUMDI003M5	D16965 Human HepG2
32	72.2	0.2	108	9	HUMDI003M5	D16965 Human HepG2
33	71	0.2	104	9	HUMALCE272	M87899 Human carci
34	71	0.2	107	11	HSU67806	U67806 Human small
35	70.6	0.2	101	10	S79560	S79560 HRX (intr
36	69.8	0.2	108	11	HSU67808	U67808 Human small
37	69.2	0.2	91	13	HUMUT8164A	L30244 Human STS U
38	69.2	0.2	99	13	HUMUT7692A	L30306 Human STS U
39	69.2	0.2	100	13	HUMUT931A	L31299 Human STS U
40	69.4	0.2	108	13	G43535	G43535 WIAP-2393-S
41	68.8	0.2	80	9	HUMBRKFAE	M36135 Human alpha
42	68.8	0.2	108	13	G32614	G32614 A009K21 Hum
43	67.6	0.2	91	13	HUMUT8164A	L30244 Human STS U
44	67.8	0.2	100	9	HUMGALNSA	D45223 Human GALNS
45	67.6	0.2	107	13	G32919	G32919 A009W27 Hum

ALIGNMENTS

```

RESULT 1
G31304          96 bp  DNA      STS      29-SEP-1998
LOCUS          sy899g1-19 Human (A.Gnirke) Homo sapiens STS genomic, sequence
DEFINITION     tagged site.
ACCESSION      G31304
VERSION        G31304.1  GI:1871333
KEYWORDS       STS.
SOURCE         human.
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
REFERENCE      1 (bases 1 to 96)
AUTHORS        Lauer, P., Meyer, N.C., Prass, C.E., Starnes, S.M., Wolff, R.K. and
                Gnirke, A.
TITLE          Clone-contig and STS maps of the hereditary hemochromatosis region
                on human chromosome 6p21.3-6p22
JOURNAL        Genome Res. 7 (5), 457-470 (1997)
MEDLINE        97294058
COMMENT        GDB: 5584195
                GDB_DSEG: D6S2377
                Contact: Andreas Gnirke
                Mercator Genetics, Inc.
                4040 Campbell Ave, Menlo Park CA, 94025, USA
                Email: gnirke@mercator.com
                Primer A: GTCCCAAGAGATATTAATGAG
                Primer B: AGGCACAGTGGGAG
                STS size: 77
PCR Profile:   Denaturation: 92 degrees C for 20 seconds
                Annealing: 60 degrees C for 45 seconds
                Polymoriation: 72 degrees C for 60 seconds
                PCR Cycles: 35
                Thermal Cycler: MJ Research PTC-200
Protocol:      Template: 30-200 ng
                Primer: each 0.8 uM
                dNTPs: each 200 uM
                Taq Polymerase: 0.05 units/ul
                Total Vol: 12 ul
Buffer:        MgCl2: 1.5 mM.
                KCl: 50 mM
                Tris-HCl: 10 mM
                pH: 8.3
                gelatin: 0.001% (w/v).
                Location/Qualifiers
                1..96
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="Human (A.Gnirke)"
                9..85
                primer_bind 9..30
                primer_bind complement(70..85)
                22 a 29 c 23 g 22 t
BASE COUNT    22 a 29 c 23 g 22 t
ORIGIN
STG
primer_bind 9..30
primer_bind complement(70..85)
22 a 29 c 23 g 22 t
BASE COUNT    22 a 29 c 23 g 22 t
ORIGIN
Query Match 0.3%; Score 94.4; DB 13; Length 96;
Best Local Similarity 99.0%; Pred. No. 4.4e-08;
Matches 95; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12871 CTGCTGGTGTCCCAAGATATAATGAGAAAATGTTCCCATGGATGCCAGATCCCC 12930
|||||
Db 1 CTGCTGGTGTCCCAAGATATAATGAGAAAATGTTCCCATGGATGCCAGATCCCC 60

QY 12931 TCTGCCCTCTTCCACATGTCCTGGGGGACAGGT 12966
|||||
Db 61 TCTGCCCTCTTCCACATGTCCTGGGGGACAGGT 96

RESULT 2
HSLDLRN2/c     108 bp  DNA      PRI      20-MAY-1992
LOCUS          Human LDL-receptor gene intron 14 fragment (normal gene).
DEFINITION     X05250
ACCESSION      X05250
VERSION        X05250.1  GI:34337
KEYWORDS       Alu repetitive sequence; low density lipoprotein receptor.
SOURCE         human.
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
REFERENCE      1 (bases 1 to 108)
AUTHORS        Horsthemke, B., Beisiegel, U., Dunning, A., Hovinga, J.R.,
                Williamson, R. and Humphries, S.
TITLE          Unequal crossing-over between two alu-repetitive DNA sequences in
                the low-density-lipoprotein-receptor gene. A possible mechanism for
                the defect in a patient with familial hypercholesterolaemia
JOURNAL        Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE        87161901
COMMENT        See X05252 for deletion junction
                Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES       Location/Qualifiers
                source
                1..108
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                1..108
                intron /note="intron XIV fragment"
                28 a 23 c 39 g 18 t
BASE COUNT    28 a 23 c 39 g 18 t
ORIGIN
Query Match 0.3%; Score 91.6; DB 10; Length 108;
Best Local Similarity 91.5%; Pred. No. 1.5e-07;
Matches 97; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 10790 CTCGGCTCACTGCAAGCTCTGCTCTCTGGTTTCATGCGCATTCCTCGCTCAGCTTCCG 10849
|||||
Db 108 CTCGGCTCACTGCAAGCTCTGCTCTCTGGTTTCATGCGCATTCCTCGCTCAGCTTCCG 49

QY 10850 AGTAGCTGGGACTACAGCGCTGCTGCCACCGCCAGCTTAATTTT 10895
|||||
Db 48 AGTAGCTGGGATTACAGCACCTGCTGCCACCGCCAGCTTAATTTT 3

RESULT 3
HSLDLRN2       108 bp  DNA      PRI      20-MAY-1992
LOCUS          Human LDL-receptor gene intron 14 fragment (normal gene).
DEFINITION     X05250
ACCESSION      X05250
VERSION        X05250.1  GI:34337
KEYWORDS       Alu repetitive sequence; low density lipoprotein receptor.
SOURCE         human.
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
REFERENCE      1 (bases 1 to 108)
AUTHORS        Horsthemke, B., Beisiegel, U., Dunning, A., Hovinga, J.R.,
                Williamson, R. and Humphries, S.
TITLE          Unequal crossing-over between two alu-repetitive DNA sequences in
                the low-density-lipoprotein-receptor gene. A possible mechanism for
                the defect in a patient with familial hypercholesterolaemia
JOURNAL        Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE        87161901
COMMENT        See X05252 for deletion junction
                Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES       Location/Qualifiers
                source
                1..108
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                1..108
                intron /note="intron XIV fragment"
                28 a 23 c 39 g 18 t
BASE COUNT    28 a 23 c 39 g 18 t
ORIGIN

```



```

HSLDLI12      108 bp      DNA      PRI      20-MAY-1992
LOCUS          Human LDL-receptor gene intron 12 fragment (normal gene) LDL - low
DEFINITION     density lipoprotein.
ACCESSION      X05248
VERSION        X05248.1 GI:34334
KEYWORDS       Alu repetitive sequence; low density lipoprotein receptor;
               repetitive sequence.
SOURCE         human
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Chordata; Vertebrata; Mammalia; Eutheria;
               Primates; Catarrhini; Hominiidae; Homo.
AUTHORS       Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
               Williamson,R. and Humphries,S.
TITLE         Unequal crossing-over between two alu-repetitive DNA sequences in
               the low-density-lipoprotein-receptor gene. A possible mechanism for
               the defect in a patient with familial hypercholesterolaemia
               Eur. J. Biochem. 164 (1), 77-81 (1987)
JOURNAL        87161901
MEDLINE
COMMENT        see X05249 for deletion junction
               Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES       location/Qualifiers
               1..108
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               complement(<1..65)
               /note="Alu repeat"
               1..108
               /note="intron XII fragment"
BASE COUNT     21 a      38 c      20 g      29 t
ORIGIN
1
...
Query Match      0.3%; Score 81.4; DB 10; Length 108;
Best Local Similarity 85.0%; Pred. No. 1.2e-05;
Matches 91; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 5461 TCGGCTCACCGCAACCTCTACTCCAGGTTCAAGCAATTCCTCGCTCAGCCTCCCGA 5520
|||||
Db 2 TCGCCTACCAACCACTCTGCTCTGGTTCAACCACTTCCTCGCTCAGCCTCCCTTA 61
|||||

Qy 5521 GTAGCTGGGATTACAGGCATCATCCACGCCACCGCCAGCTAATTTGTA 5567
|||||
Db 62 GTAGCTGGGATTACAGCATGTGCCACCAACGCCCGCGCTGATTGTGA 108
|||||

RESULT 14
HSU67808/c
LOCUS          Human small cytoplasmic Alu transcript.
DEFINITION     U67808
ACCESSION      U67808
VERSION        U67808.1 GI:2289922
KEYWORDS       Alu.
SOURCE         human.
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
               Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS       Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE         cDNAs derived from primary and small cytoplasmic Alu (scAlu)
               transcripts
               J. Mol. Biol. 271 (2), 222-234 (1997)
JOURNAL        97415756
MEDLINE
REFERENCE      2 (bases 1 to 108)
AUTHORS       Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE         Direct Submission
JOURNAL        Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
               Children's Hospital of Philadelphia, 1004F Abramson Research
               Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES       location/Qualifiers
               1..108
               /organism="Homo sapiens"

```

```

/db_xref="taxon:9606"
/clone="TscAlu7"
repeat_region 1..108
/note="scAlu"
/rpt_family="Alu"
/rpt_type="dispersed"
BASE COUNT     22 a      37 c      28 g      21 t
ORIGIN
1
...
Query Match      0.3%; Score 79; DB 11; Length 108;
Best Local Similarity 89.5%; Pred. No. 3.5e-05;
Matches 85; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 10908 AGAGATGGGGTTTACCACATGTAGCCAGGATGGTCTCATCTCTGACCTCGTGATCCAC 10967
|||||
Db 95 AAGACGGAGGTTTACCACATGTGGCCAGGCTGGTCTCAAACTCTGACCTTGTGTATCCAC 36
|||||

Qy 10968 CCGCTTTGGCCTCCCAAGTGTCTGGGATTACAGGC 11002
|||||
Db 35 CCGACTTGGCCTCCCAAGTGTCTGGGATTACAGGC 1

RESULT 15
HUMALCE43/c
LOCUS          HUMALCE43 110 bp ss-RNA PRI 15-APR-1994
DEFINITION     Human carcinoma cell-derived Alu RNA transcript, clone CE43.
ACCESSION      M87900
VERSION        M87900.1 GI:174876
KEYWORDS       Alu repeat.
SOURCE         Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
               Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      1 (bases 1 to 110)
AUTHORS       Smet,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE         Alu RNA transcripts in human embryonal carcinoma cells. Model of
               post-transcriptional selection of master sequences
               J. Mol. Biol. (1992) In press
JOURNAL        Location/Qualifiers
FEATURES       1..110
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /cell_line="NTERA2D1"
               /dev_stage="embryo"
               /sex="male"
               /tissue_type="carcinoma"
BASE COUNT     27 a      31 c      34 g      18 t
ORIGIN
1
...
Query Match      0.3%; Score 78.8; DB 9; Length 110;
Best Local Similarity 84.0%; Pred. No. 3.8e-05;
Matches 89; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 10917 GTTTCACCATGTAGCCAGGATGGTCTCATCTCTGACCTCGTGATCCACCGCTTTGG 10976
|||||
Db 110 GTTTCGTCATGTAGCCAGGCTGGTCTTGAATACTAGCTCGCAATCTCTCGCTTGG 51
|||||

Qy 10977 CTCTCCAAAGTCTGGGATTACAGCGGTGAGCCACCGTGCCCGGCC 11022
|||||
Db 50 CTCTCCAAAGTCCGGGATTGTAGGTGTGAGCCACCGCCCGCGGCC 5

Search completed: June 15, 2000, 04:50:04
Job time: 58288 sec

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query		Length	DB	ID	Description
		Match					
C 1	70	0.2	0.2	108	1	X12095	Human biallelic po
C 2	67.2	0.2	100	1	X24892		Human gene signatu
C 3	66	0.2	103	1	T20927		Human gene signatu
C 4	65.2	0.2	108	1	X12095		Human biallelic po
C 5	64	0.2	100	1	X12086		Human biallelic po
C 6	63.6	0.2	100	1	X12087		Human biallelic po
C 7	63.6	0.2	100	1	X12085		Human biallelic po
C 8	62.4	0.2	100	1	T24892		Human gene signatu
C 9	62.6	0.2	103	1	T26213		Human gene signatu
C 10	62.2	0.2	108	1	T26828		Human gene signatu
C 11	60.2	0.2	103	1	T20927		Human gene signatu
C 12	59.4	0.2	91	1	T25854		Human gene signatu
C 13	59.4	0.2	108	1	T26828		Human gene signatu
C 14	58.6	0.2	108	1	T25009		Human gene signatu
C 15	57.8	0.2	108	1	T25009		Human gene signatu
C 16	57.2	0.2	91	1	T25854		Human gene signatu
C 17	55.6	0.2	103	1	T26213		Human gene signatu
C 18	55.2	0.2	87	1	T21566		Human gene signatu
C 19	55.4	0.2	110	1	T25260		Human gene signatu
C 20	55	0.2	95	1	T22572		Human gene signatu
C 21	54.2	0.2	93	1	T23131		Human gene signatu
C 22	54	0.2	99	1	T20931		Human gene signatu
C 23	53.4	0.2	87	1	T21566		Human gene signatu
C 24	53.2	0.2	93	1	T22572		Human gene signatu
C 25	53.4	0.2	97	1	T26728		Human gene signatu
C 26	53	0.2	109	1	T23895		Human gene signatu
C 27	52.4	0.2	70	1	N60231		Normal chromosome
C 28	52.6	0.2	93	1	T24259		Human gene signatu
C 29	52.2	0.2	100	1	X12085		Human biallelic po
C 30	52.4	0.2	110	1	T25260		Human gene signatu
C 31	52	0.2	69	1	T29016		Probe to internal
C 32	52	0.2	84	1	T25848		Human gene signatu
C 33	52	0.2	93	1	T25688		Human gene signatu
C 34	52	0.2	93	1	T25688		Human gene signatu

CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases. 23 C; 28 G; 37 T;
SQ Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;

Query Match 0.2%; Score 65.2; DB 1; Length 108;
Best Local Similarity 80.6%; Pred. No. 0.082;
Matches 87; Conservative 1; Mismatches 19; Indels 1; Gaps 1;

QY 21852 TGTATTTTGTAGACGGGTTTCCACATGTTGGTCAGGCTGTCGGAACTCCTGAC 21911
Db 1 TTTCTTTTGTAGAGATGAGGTTTTCCTRTGTGGCCAGGATGTCGGAACCTCTGAC 60

QY 21912 CTCAGTGATGTCGCCACTCAGCTCC-AAAAGTCGTGGGATTACAG 21958
Db 61 TTCAGTGATGCTGCTGCTGGCCCTCCAAAGATGCTGGGATTATAG 108

RESULT 5
X12086/c
ID X12086 standard; DNA; 100 BP.
AC X12086;
DE 30-MAR-1999 (first entry)
KW Human biallelic polymorphic DNA fragment EST98276b.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match 0.2%; Score 64; DB 1; Length 100;
Best Local Similarity 77.6%; Pred. No. 0.12;
Matches 76; Conservative 1; Mismatches 21; Indels 0; Gaps 0;

QY 4303 GTGGCTCAGGCTGTAAATCCCAAGCACTTCAGGAGCTGAGTTGGGAGATCCTTGAGCT 4362
Db 99 GTGACTCACACCTATATCTGGCACTTTAGGAGGCTTAGGAGGAGGATGTTTGAAC 40

QY 4363 CAGGAGTTCAAGACCACTTTTGGGCAACATAGCAAGTCT 4400
Db 99 GTGACTCACACCTATATCTGGCACTTTAGGAGGCTTAGGAGGAGGATGTTTGAAC 40

Db 39 CAGGAGCTCAAGACCATCTCTGGGAAACATAGCAAGACT 2

RESULT 6
X12087/c
ID X12087 standard; DNA; 100 BP.
AC X12087;
DE 30-MAR-1999 (first entry)
KW Human biallelic polymorphic DNA fragment EST98276a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match 0.2%; Score 63.6; DB 1; Length 100;
Best Local Similarity 78.1%; Pred. No. 0.14;
Matches 75; Conservative 1; Mismatches 20; Indels 0; Gaps 0;

QY 22822 GAGGCTCTGCTGTAAATCCCAAGCACTTCAGGAGCTGAGTTGGGAGGAGGATGCTTGAGCC 22881
Db 99 GTGACTCACACCTATATCTGGCACTTTAGGAGGCTTAGGAGGAGGATGTTTGAAC 40

QY 22882 CAGGAATCAAGACCAAGCTCTGGGAAACATAGGAGAG 22917
Db 39 CAGGAGCTCAAGACCAAGCACTCTGGGAAACATAGCAAGA 4

RESULT 7
X12085/c
ID X12085 standard; DNA; 100 BP.
AC X12085;
DE 30-MAR-1999 (first entry)
KW Human biallelic polymorphic DNA fragment EST98276c.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.

QY 15835 TCTCAGCTCTACTGCAACCTGCACCTCTGCTGGTTCAAGGGAT 15875
 II |IIIIIIII III |IIIIIIII IIIII IIII
 Db 42 TCATAGCTCACTGTAACCAACCAACTCTCGGACTCAAGTGAT 2

RESULT 10
 T26828
 ID T26828 standard; cDNA to mRNA; 108 BP.
 AC T26828;
 DT 14-NOV-1996 (first entry)
 DE Human gene signature HUMGS09078.
 KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 human; cloning; mapping; non-biased library; diagnosis; detection;
 cell typing; abnormal cell function; ss.
 KW Homo sapiens.
 PS WO9514772-A1.
 PN 01-JUN-1995.
 PD 11-NOV-1994; J01916.
 PF 12-NOV-1993; JP-355504.
 PR (MATS/) MATSUBARA K.
 PA (OKUB/) OKUBO K.
 PI Matsubara K, Okubo K;
 DR WPI; 95-206931/27.
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function, by preparing cDNA that
 PT reflects relative abundance of corresp. mRNA in specific human
 PT tissues
 PS Claim 1; Page 2182; 2245pp; Japanese.
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC given in T19001-T26837 and which is able to hybridise to part of
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined (esp. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.
 SQ Sequence 108 BP; 18 A; 33 C; 23 G; 28 T;

Query Match 0.28; Score 62.2; DB 1; Length 108;
 Best Local Similarity 75.2; Pred. No. 0.23;
 Matches 76; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 10945 GATCTCTGACCTCGTGATCCACCGCTTTGGCTCCCAAGTCTGGGANTACAGCGT 11004
 II |IIIIIIIIIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII
 Db 1 GATCTCTGACCTCGTGATCCCGCGGTTTCGGCTCCCATAGTCTGGGNTTACAGGCAT 60

QY 11005 GAGCCACCGTCCCGCGCTACTTCACTTCTTCATTTAAAA 11045
 II |IIIIII IIIII III |IIIIII IIIII IIIII IIIII IIIII
 Db 61 GAGCCACCGTCCCGCGCTGTTTATTTCTTATACTGTACA 101

RESULT 11
 T20927
 ID T20927 standard; cDNA to mRNA; 103 BP.
 AC T20927;
 DT 24-JUL-1996 (first entry)
 DE Human gene signature HUMGS02180.
 KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 human; cloning; mapping; non-biased library; diagnosis; detection;
 cell typing; abnormal cell function; ss.
 KW Homo sapiens.
 PS WO9514772-A1.
 PN 01-JUN-1995.
 PD 11-NOV-1994; J01916.

CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 91 BP; 18 A; 22 C; 28 G; 18 T;

Query Match 0.2%; Score 59.4; DB 1; Length 91;
Best Local Similarity 77.5%; Pred. No. 0.6; Mismatches 0; Gaps 0;
Matches 69; Conservative 0;

Qy 15787 TGAGACAGGCTCTCACTATCACCAGGCTGGAGTGGCAGCAGTCTCAGCTCACT 15846
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 90 TGAGACAGNNCTCAGCTGTCCACCNAGGCTGGAGCGCAGAGTGCCATCTCAGCTCACT 31

Qy 15847 GCACCTGTCACCTCTCGTGGTTCAGGGAT 15875
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 30 TGAACNCCTGCCTCCTAGGCTCAAGTGAT 2

RESULT 13
T26828/c
ID T26828 standard; cDNA to mRNA; 108 BP.
AC T26828;
DT 14-NOV-1996 (first entry)
DE Human gene signature HUMGS09078.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 2182; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 18 A; 23 C; 28 G; 28 T;

Query Match 0.2%; Score 59.4; DB 1; Length 108;
Best Local Similarity 83.5%; Pred. No. 0.6; Mismatches 0; Gaps 0;
Matches 66; Conservative 0;

Qy 6119 CAGCGGGTGCAGTGGCTCATCTGTAATCCAGCACTTTGGGAGTTCGAGCGGCTG 6178
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 79 CAGCGGGCGTGGTGGCTCATCTGTAATCCAGCACTATGGGAGCGGACGAGCGG 20
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 6179 ATCAGAGGTCAGGATTC 6197
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 19 ATGACGAGGTCAGGAGATC 1

RESULT 14
T25009/c
ID T25009 standard; cDNA to mRNA; 108 BP.
AC T25009;
DT 07-NOV-1996 (first entry)
DE Human gene signature HUMGS07131.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1748; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 34 A; 31 C; 26 G; 15 T;

Query Match 0.2%; Score 58.6; DB 1; Length 108;
Best Local Similarity 71.0%; Pred. No. 0.79;
Matches 76; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 15769 TTTTNTTTTTTTTTTTTGTGAGACAGAGTCTCTATCATCACCGAGTGGAGTG 15828
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108 TTTGNTGTTGTTGTTGTTTTCACACAGGCTCTGCTCTACTCAGGCTGGATNCAGTG 49

Qy 15829 GCACAACTCAGCTCACTGCACTGCACCTGCCTCTGGTTCAGGGAT 15875
|| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 48 GCGTGACCATGGCTCACTGCGAGCCTTGGCCTCATGGGCTCAGGCGAT 2

RESULT 15
T25009
ID T25009 standard; cDNA to mRNA; 108 BP.
AC T25009;
DT 07-NOV-1996 (first entry)
DE Human gene signature HUMGS07131.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2000, 15:22:36 ; Search time 8512.35 Seconds
(without alignments)
13809.057 Million cell updates/sec

Title: US-08-852-495C-1_COPY_28000_57000
Perfect score: 29001
Sequence: 1 ATATCACACAAAACACACAT.....TTAGCAGCACACAGGTAGGGT 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues
Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : EST:
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_gss1:*
82: gb_gss2:*
83: gb_gss3:*
84: gb_gss4:*
85: em_gss1:*
86: em_gss2:*
87: em_gss3:*
88: em_gss4:*
89: gb_gss5:*
90: gb_gss6:*
91: gb_gss7:*
92: gb_gss8:*
93: gb_gss9:*
94: gb_gss10:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	93	0.3	109	30	AA243009	AA243009 zr25h02.s
2	93	0.3	109	84	B17434	B17434 345K2.TVB C
3	92.2	0.3	105	61	A1832832	A1832832 at72909.x
4	91.6	0.3	106	37	AA703692	AA703692 ag81a10.r
5	91.6	0.3	106	105	AQ264176	AQ264176 CITBI-El-
6	91.4	0.3	109	94	AQ028426	AQ028426 CIT-HSP-2
7	90.2	0.3	103	108	AQ535244	AQ535244 RPT-11-3
8	89.4	0.3	107	35	AA565533	AA565533 nk42b11.s
9	89	0.3	105	28	AA078003	AA078003 7H12D08 C
10	89	0.3	105	61	A1832832	A1832832 at72909.x
11	88.6	0.3	103	84	B48914	B48914 RPT111-4A12
12	88.6	0.3	103	108	AQ535244	AQ535244 RPT-11-3
13	88.2	0.3	109	105	AQ265749	AQ265749 CITBI-El-
14	87.6	0.3	102	94	AQ004934	AQ004934 CIT-HSP-2
15	87.6	0.3	103	38	AA807640	AA807640 nx08b05.s
16	87.6	0.3	110	64	AW083640	AW083640 xc49f02.x
17	87.2	0.3	110	33	AA442529	AA442529 zv68b02.r
18	86.6	0.3	102	30	AA226656	AA226656 nc19f09.s
19	86.8	0.3	106	94	AQ062963	AQ062963 CIT-HSP-2
20	86.8	0.3	110	39	AA897366	AA897366 am06h02.s
21	86.6	0.3	110	94	AQ003188	AQ003188 RPT111-1D
22	85.8	0.3	105	105	AQ282107	AQ282107 RPT111-94
23	85.4	0.3	103	94	AQ028649	AQ028649 CIT-HSP-2
24	85.6	0.3	110	33	AA442529	AA442529 zv68b02.r
25	85	0.3	101	39	AA835205	AA835205 ak64h01.s
26	85.2	0.3	106	38	AA812141	AA812141 OB48h02.s
27	84.8	0.3	104	105	AQ231855	AQ231855 RPT111-11
28	85	0.3	110	64	AW083640	AW083640 xc49f02.x
29	84.6	0.3	100	35	AA564832	AA564832 nj22a06.s
30	84.4	0.3	102	36	AA654562	AA654562 nt75f10.s
31	84.4	0.3	102	84	B48088	B48088 RPT111-4N6
32	84.6	0.3	103	108	AQ584425	AQ584425 RPT-11-4
33	84.6	0.3	108	84	B65160	B65160 CIT-HSP-201
34	84.2	0.3	107	24	H67040	H67040 yu68c01.r1
35	84.4	0.3	110	79	AW250394	AW250394 2822460.3
36	84.4	0.3	110	79	AW250394	AW250394 2822460.3
37	83.6	0.3	100	35	AA564832	AA564832 nj22a06.s
38	83.8	0.3	103	108	AQ534922	AQ534922 RPT-11-3
39	83.8	0.3	105	109	AQ373292	AQ373292 RPT-11-4
40	83.2	0.3	98	24	H67549	H67549 yu68f10.s1
41	83.4	0.3	103	107	AQ485214	AQ485214 RPT-11-2
42	83	0.3	101	33	AA381369	AA381369 EST94442
43	83.2	0.3	104	108	AQ544583	AQ544583 CITBI-El-
44	83.2	0.3	107	33	AA385808	AA385808 EST99495
45	83.2	0.3	107	103	AQ240182	AQ240182 CIT-HSP-2

ALIGNMENTS

RESULT 1

AA243009

LOCUS

DEFINITION

zr25h02.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone IMAGE:664467 3' similar to contains Alu repetitive element; contains element LTR1 repetitive element ;, mRNA sequence.

ACCESSION

AA243009

VERSION

AA243009.1 GI:1873869

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 109)

Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Willson, R.

TITLE

WashU-NCI human EST Project

JOURNAL

Unpublished (1997)

COMMENT

On Dec 3, 1996 this sequence version replaced gi:1126869.

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1127 Std Error: 0.00
Seq primer: -41ml3 fwd. ET from Amersham
High quality sequence stop: 102.

FEATURES

Location/Qualifiers

1..109

/organism="Homo sapiens"

/db_xref="GDB:5426481"

/db_xref="taxon:9606"

/clone="IMAGE:664467"

/clone_lib="Stratagene NT2 neuronal precursor 937230"

/tissue_type="neuroepithelial cells"

/dev_stage="Ntera-2 neuroepithelial cells"

/lab_host="SOLR (kanamycin resistant)"

/note="Organ: brain; Vector: pBluescript SK-; Site:1: EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Uninduced, exponentially growing neuroepithelial cells (Ntera-2/cl.D1). Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTCCTTTT 3'"

BASE COUNT

19 a 30 c 30 g 30 t

ORIGIN

Query Match 0.3%; Score 93; DB 30; Length 109;

Best Local Similarity 90.8%; Pred. No. 0.22; Indels 0; Gaps 0;

Matches 99; Conservative 0; Mismatches 10;

QY 10897 GTATTTTATTAGATGGGGTTTCCACCATTTAGCCAGGATGCTCGATCTCCTGACC 10956

Db 1 GTATTTTATTAGATGGGGTTTCCACCATTTAGCCAGGATGCTCGATCTCCTGACC 60

QY 10957 TCGTGATCCACCCCTTTGGCCCTCCCAAGTCTGGGATTACAGCGGTG 11005

Db 61 TCGTGATCCCGCCCTCGCGCTCCCAAGTCTGGGATTACAGCGGTG 109

RESULT 2

B17434/c

LOCUS

B17434 109 bp DNA GSS 04-JUN-1998

DEFINITION

345K2.TVB CIT978SKA1 Homo sapiens genomic clone A-345K02, genomic survey sequence.

ACCESSION

B17434

VERSION

B17434.1 GI:2125183

KEYWORDS

GSS.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 109)

Adams, M.D., Kelley, J.M., Rounsley, S.R. and Venter, J.C.

AUTHORS

Use of a BAC End Sequence Database for Sequence-Ready Map Building

TITLE

Unpublished (1997)

JOURNAL

Other_GSSs: 345K02.TP 345K02.TPB

COMMENT

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdamadset@igrr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html

Seq primer: T7

Class: BAC ends.

FEATURES
source

Location/Qualifiers
1. .109
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="A-345K02"
/clone_lib="CIT978SKA1"
/sex="Female"
/cell_type="Fibroblast"
/note="vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
CalTech Human BAC Library A1"
24 a 30 c 31 g 24 t

BASE COUNT
ORIGIN

Query Match 0.3%; Score 93; DB 84; Length 109;
Best Local Similarity 90.8%; Pred. No. 0.22; Indels 0; Gaps 0;
Matches 99; Conservative 0; Mismatches 10;

Oy 21860 TAGTAGACGGGGTTTCACCATGTTGTGTCAGGCTGTGGAACCTCTGACCTCAGGTG 21919
|||||
Db 109 TAGTTGACGGGGTTTCACCATGTTGTGTCAGGCTGTGGAACCTCTGACCTCAGGTG 50
|||||
Oy 21920 ATCTGCCCACTCAGCTCCCAAGTCTGGGATACAGGATGAGCCA 21968
|||||
Db 49 ATCCGCCACATCAGCTCCCAAGTCTGGGATACAGGATGAGCCA 1
|||||

RESULT 3

AA1832832 105 bp mRNA EST 13-JUL-1999
LOCUS at72909.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone
DEFINITION IMAGE:2377600 3' similar to contains Alu repetitive
element; contains element MER22 repetitive element ;, mRNA sequence.
A1832832
ACCESSION A1832832.1 GI:5454812
VERSION EST.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 105)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)

JOURNAL
COMMENT On Dec 20, 1995 this sequence version replaced gi:1133644.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40up from Gibco.
Location/Qualifiers

FEATURES
source

1. .105
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2377600"
/clone_lib="Barstead colon HPLRB7"
/sex="male"
/dev_stage="adult, age 25"
/lab_host="DH10B (phage resistant)"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTACGAATCTGAAGTGGAGCGGCCCTTTTTTTTTTTTTTTTTTTTTTTT
3']; double stranded cDNA was ligated to Eco RI adaptors
[5' AATCTACTAGTAAT 3' and 5' ATTACTAGTG 3'], digested
with Not I and cloned into the Not I and Eco RI sites of

the modified pT73 vector. Library constructed by Bob
Barstead." 17 a 35 c 27 g 26 t

Query Match 0.3%; Score 92.2; DB 61; Length 105;
Best Local Similarity 92.4%; Pred. No. 0.27;
Matches 97; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 5411 GAGATGGAGTTTCGCTCTTTGTCGCCAGCTGGAGTCAATGGCGGTCTACC 5470
|||||
Db 1 GAGACAGAGTTTCGCTCTTTGTCGCCAGCTGGAGTCAATGGCGGTCTACC 60
|||||

Oy 5471 GCACCTCTACCTCCAGGTTCAACGAATTCCTCCGCTCAGCCT 5515
|||||
Db 61 GCACCTCCACCTCCCGGTTCAACGGAATTCCTCCGCTCAGCCT 105
|||||

RESULT 4

AA703592 106 bp mRNA EST 24-DEC-1997
LOCUS ag81a10.r1 StrataGene hNT neuron (#937233) Homo sapiens cDNA clone
DEFINITION IMAGE:1140858 5' similar to contains Alu repetitive element; , mRNA
sequence.
A703592
ACCESSION AA703592.1 GI:2713610
VERSION EST.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 106)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)

JOURNAL
COMMENT On Sep 12, 1996 this sequence version replaced gi:11397630.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28ml3 revl ET from Amersham
High quality sequence stop: 53.
Location/Qualifiers

1. .106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1140858"
/clone_lib="StrataGene hNT neuron (#937233)"
/dev_stage="hNT neurons"
/lab_host="SOLR (kanamycin resistant)"
/note="vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT.
Differentiated, post mitotic hNT neurons. Average insert
size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'
GAATTCGGCAGGAG 3' -3' adaptor sequence: 5'
CTCAGATTTTTTTTTTTTTTTTTTTT 3'.

FEATURES
source

19 a 29 c 29 g 29 t

BASE COUNT
ORIGIN

Query Match 0.3%; Score 91.6; DB 37; Length 106;
Best Local Similarity 91.5%; Pred. No. 0.32;
Matches 97; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Oy 10900 TTTTATTAGATGGGGTTTCACCATGTTACCCAGGATGGTCTCGATCTCGACCTCG 10959

VERSION	AQ028426.1	GI:3268648
KEYWORDS	GSS.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	1 (bases 1 to 109); Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.	
TITLE	Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)	
JOURNAL	Unpublished (1998)	
COMMENT	Contact: Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: mdadams@tigr.org Clones are available from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html . Seq primer: M13-21 Class: BAC ends.	
FEATURES	Location/Qualifiers	
source	1..109	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="2313G15"	
	/clone_lib="citr-HSP"	
	/sex="Male"	
	/cell_type="Sperm"	
	/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"	
BASE COUNT	19 a 36 c 25 g 29 t	
ORIGIN		
Query Match	0.3%; Score 91.4; DB 94; Length 109;	
Best Local Similarity	89.9%; Pred. No. 0.33; 11; Indels 0; Gaps 0;	
Matches	98; Conservative 0; Mismatches 11; Indels 0; Gaps 0;	
QY 15779	TTTTTTTTTGAGACAGAGTCTCACTCTATCACCAGGCTGGAGTGCAGTGCACAAATCTC 15838	
Db 1	TTGTTTTCTGAGGGGACTCTCACTCTGTCAACCAGGCTGGAGTGCAGTGCACAGTCTG 60	
QY 15839	AGCTCACTGCAACCTCACCTCTCGGGTTCAAGGGATTCTCTACCTAA 15887	
Db 61	AGCTCACTGCAACCTCACCTCTCGGGTTCAAGGGATTCTCTCTGCTCA 109	
RESULT 7		
AQ535244	103 bp DNA	GSS 18-MAY-1999
LOCUS	RPCI-11-317H22.TV RPCI-11 Homo sapiens genomic clone	
DEFINITION	RPCI-11-317H22, genomic survey sequence.	
ACCESSION	AQ535244	
VERSION	AQ535244.1	GI:4846934
KEYWORDS	GSS.	
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 103)	
AUTHORS	Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.	
TITLE	Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building	
JOURNAL	Unpublished (1997)	
COMMENT	Contact: Shaying Zhao, William Nierman, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research	

genomic DNA (cosmids). The resulting direct-selected cDNA was cloned into a plasmid vector using a non-directional uracil DNA glycosylase (UDG)-mediated cloning strategy."

BASE COUNT 20 a 33 c 23 g 29 t
ORIGIN

Query Match 0.3%; Score 89; DB 28; Length 105;
Best Local Similarity 90.5%; Pred. No. 0.61;
Matches 95; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 9922 TACAAAAATTACCGGGGCGGTGGTGCAGCGCTCTAATCCAGCTACTTGGAGGCTGA 9981

Db 105 TACAAAAATTAGCTGGCGATGTAGCGACGCATGTAATCCAGCTACTTGGAGGCTGA 46

QY 9982 GACAGGAGNATCGTTGAACCTGGGAGGAGGTTACAGTGAGC 10026

Db 45 GACCCGAGAAATGCTTGACACCGAGGAGGAGGTTGCAGTGAGC 1

RESULT 10
AI832832/c

LOCUS AI832832 105 bp mRNA EST 13-JUL-1999

DEFINITION at72g09.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone IMAGE:2377600 3' similar to contains Alu repetitive element; contains element MER22 repetitive element ; , mRNA sequence.

ACCESSION AI832832

VERSION AI832832.1 GI:5454812

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 105)

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

WashU-NCI human EST Project

Unpublished (1997)

On Dec 20, 1995 this sequence version replaced gi:1133644.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40UP from Gibco.

Location/Qualifiers

FEATURES

Source

1..105

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2377600"

/clone_lib="Barstead colon HPLRB7"

/sex="male"

/dev_stage="adult, age 25"

/lab_host="DH10B (phage resistant)"

/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACGAATCTGAAGTGGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [5' AATCTACTAGTAT 3' and 5' ATTACTAGT 3'], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed by Bob Barstead."

BASE COUNT 17 a 35 c 27 g 26 t

ORIGIN

Query Match 0.3%; Score 89; DB 61; Length 105;

Best Local Similarity 90.5%; Pred. No. 0.61;
Matches 95; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 20776 AAGCTGACACAGGAAATCGCTTGAACCTGGGAGCGGAGGTGTGTGAGCCGAGATCA 20835

Db 105 AGGCTGAGGCGAGGAGATCGCTTGAACCCGGGAGGTGCGGTGAGCCAAGATCG 46

QY 20836 TGCCATTGCACTCCAGCTGGGCAACAAGAGCGAAATCCGTC 20880

Db 45 CACCAATGCACTCCAGCTGGGCAACAAGAGCGAAATCTGTCTC 1

RESULT 11

LOCUS B48914/c

DEFINITION RPC111-4A12.TP RPC1-11 Homo sapiens genomic clone RPC1-11-4A12, genomic survey sequence.

ACCESSION B48914

VERSION B48914

KEYWORDS GSS.

SOURCE GSS.

human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 103)

AUTHORS Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and Venter, J.C.

Use of BAC End Sequences for Sequence-Ready Map Building

Unpublished (1997)

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are derived from the human BAC library RPC1-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1..103

/organism="Homo sapiens"

/db_xref="GDB:7501163"

/db_xref="taxon:9606"

/clone="RPC1-11-4A12"

/clone_lib="RPC1-11"

/sex="Male"

/cell_type="Lymphocytes"

/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;

RPC111 Human Male BAC Library"

BASE COUNT 30 a 28 c 30 g 15 t

ORIGIN

Query Match 0.3%; Score 88.6; DB 84; Length 103;

Best Local Similarity 91.3%; Pred. No. 0.68;

Matches 94; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 10896 TGTATTTTATAGATGGGGTTTCACCATGTTAGCCAGGATGTCCTCGATCTCTGAC 10955

Db 103 TGTATTTTATAGAGCGGGTTTCACCGTTTATAGCCGGATGTCCTCGATCTCTGAC 44

QY 10956 CTCGTGATCCACCGCTTTGGCCTCCCAAGTGCGGGATTAC 10998

Db 43 CTCGTGATCCGCGCGCTCGGCTCCCAAGTGCTGGGGCTTAC 1

```

RESULT 12
A0535244/c 103 bp DNA GSS 18-MAY-1999
LOCUS
DEFINITION RCI-11-317H22-TV RCI-11 Homo sapiens genomic clone
RCI-11-317H22, genomic survey sequence.
ACCESSION A0535244
VERSION A0535244.1 GI:4846934
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and
Venter, J.C.
TITLE Use of BAC End Sequences from Library RCI-11 for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetigr.org
Clones are derived from the human BAC library RCI-11. For BAC
library availability, please contact Pieter de Jong
(pleteredejong.med.bufo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.bufo.edu/ordering) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.
FEATURES
source
location/Qualifiers
1..103
/organism="Homo sapiens"
/db_xref="GDB:7621533"
/db_xref="taxon:9606"
/clone="RCI-11-317H22"
/clone_lib="RCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RCI11 Human Male BAC Library"
BASE COUNT 31 a 27 c 27 g 18 t
ORIGIN

Query Match 0.3%; Score 88.6; DB 108; Length 103;
Best Local Similarity 91.3%; Pred. No. 0.68;
Matches 94; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 21848 TTTTGTATTTTATAGACAGCGGGTTTACCATGTTGGTCAGGCTGGTGTGAACCTCC 21907
||||| ||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 103 TTTTGTATTTATAGACAGACGGGTTTACCATGTTGGCCAGGCTGGTGTGAACTCC 44
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 21908 TGACCTCAGGTGATCTGCCACCTCAGCCTCCCAAGTGTGG 21950
||||| ||||||| ||||||| || ||||||| |||||||
Db 43 TGACCTCAAGTGATCTGCCCGCTTCTGGCCTCCCAAGTGTGG 1
||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 13
A0265749/c 109 bp DNA GSS 27-OCT-1998
LOCUS
DEFINITION CITBI-E1-2510E2.TR CITBI-E1 Homo sapiens genomic clone 2510E2,
genomic survey sequence.
ACCESSION A0265749
VERSION A0265749.1 GI:3791503
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 109)
AUTHORS Adams, M.D., Rounsley, S.D., Zhao, S., Field, C.E., Bass, S., Linher, K.,
Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H.,
Simon, M. and Venter, J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org

REFERENCE 1 (bases 1 to 109)
AUTHORS Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K.,
Golden, K., Granger, D., Suh, E., Wible, C., Shizuya, H.,
Simon, M., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H.,
Simon, M. and Venter, J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org

FEATURES
source
location/Qualifiers
1..109
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2510E2"
/clone_lib="CITBI-E1"
/sex="male"
/cell_type="sperm"
/notes="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"
BASE COUNT 24 a 29 c 26 g 30 t
ORIGIN

Query Match 0.3%; Score 88.2; DB 105; Length 109;
Best Local Similarity 88.1%; Pred. No. 0.73;
Matches 96; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 9889 GCCAACATGGTGAACCTCTCTACTAAAATAACAAAATAGCCGGCATGGTGGT 9948
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 109 GCCAGCATGGTGAACCTCTCTACTAGAAAATAACAAAATAGTCGAGCGCTGGCG 50
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 9949 CACGCTCTTAATCCAGCTACTTGGAGCTGAGACAGAGAAATCGCTT 9997
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 49 CATGCTCTTAATCCAGCTACTTGGAGCTGAGCAGGAGCATCACTT 1
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 14
A0004934
LOCUS
DEFINITION CIT-HSP-2292A10.TF CIT-HSP Homo sapiens genomic clone 2292A10,
genomic survey sequence.
ACCESSION A0004934
VERSION A0004934.1 GI:3082379
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 102)
AUTHORS Adams, M.D., Rounsley, S.D., Zhao, S., Field, C.E., Bass, S., Linher, K.,
Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H.,
Simon, M. and Venter, J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org

```

Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13-21;
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1..102
/organism="Homo sapiens"
/db_xref="GDB:7151269"
/db_xref="taxon:9606"
/clone="2292A10"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"

BASE COUNT 17 a 39 c 26 g 20 t

ORIGIN

Query Match 0.3%; Score 87.6; DB 94; Length 102;
Best Local Similarity 91.2%; Pred. No. 0.87; Mismatches 0; Indels 0; Gaps 0;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5433 GCCCAGCTGGAGTGCATGGCGGATCTCGGCTCACGCAACCTCTACCTCCCGAGGTTTC 5492
|||||
Db 1 GCCCAGCTGGAGTGCATGGCGGATCTCGGCTCACGCAACCTCTACCTCCCGAGGTTTC 60
QY 5493 AAGCAATTCCTCGCTCAGCTCCCGAGTAGCTGGGATTAC 5534
|||||
Db 61 AAGCAGCTCTCTGCTTAGGCTCCCGAGTAGCTGGCATTAC 102

RESULT 15

AA807640 103 bp mRNA EST 05-MAR-1998
LOCUS nx08b05.s1 NCI-CGAP_GC3 Homo sapiens cDNA clone IMAGE:1255473 3'
DEFINITION similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION AA807640
VERSION AA807640.1 GI:2877108
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2151346.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 774 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 87.

FEATURES
source

Location/Qualifiers
1..103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1255473"
/clone_lib="NCI-CGAP_GC3"
/tissue_type="pooled germ cell tumors"

/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker: 1st strand cDNA was prepared from 3 pooled germ cell tumors, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is not normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 19 a 27 c 30 g 27 t

ORIGIN

Query Match 0.3%; Score 87.6; DB 38; Length 103;
Best Local Similarity 91.2%; Pred. No. 0.87; Mismatches 0; Indels 0; Gaps 0;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10905 ATTAGAGATGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATC 10964
|||||
Db 2 AGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATC 61
QY 10965 CACCCGGCTTGGCCTCCCAAGTCTGGGATTACAGGCGTGA 11006
|||||
Db 62 CGCTCACCTCGGCCCTCCCAAGTCTGGGATTACAGGCGTGA 103

Search completed: June 14, 2000, 23:50:22

Job time: 40389 sec

Wed Jun 21 14:43:56 2000

us-08-852-495c-1_copy_28000_57000.rni

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2000, 20:28:41 ; Search time 372.61 seconds
(without alignments)

10117.003 Million cell updates/sec

Title: US-08-852-495C-1_COPY_28000_57000

Perfect score: 29001

Sequence: 1 ATATCAACAAAACACACAT.....TTAGCAGCACAGGTAGGGT 29001

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10

Maximum DB seq length: 110

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/6_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	80	0.3	105	4	US-08-481-658B-65
C 2	80	0.3	105	4	US-08-477-504A-65
C 3	80	0.3	105	4	US-08-486-756A-65
C 4	80	0.3	105	4	US-08-485-862B-65
C 5	80	0.3	105	5	US-08-787-739-65
C 6	79.4	0.3	105	4	US-08-481-658B-65
C 7	79.4	0.3	105	4	US-08-477-504A-65
C 8	79.4	0.3	105	4	US-08-486-756A-65
C 9	79.4	0.3	105	4	US-08-485-862B-65
C 10	79.4	0.3	105	5	US-08-787-739-65
C 11	65.4	0.2	84	3	US-08-454-557C-91
C 12	65.4	0.2	84	4	US-08-340-426D-91
C 13	65.4	0.2	84	4	US-08-450-673C-91
C 14	65.4	0.2	84	6	PCT-US95-17111A-91
C 15	62.8	0.2	84	3	US-08-454-557C-91
C 16	62.8	0.2	84	4	US-08-340-426D-91
C 17	62.8	0.2	84	4	US-08-450-673C-91
C 18	62.8	0.2	84	6	PCT-US95-17111A-91
C 19	60.8	0.2	85	3	US-08-454-557C-92
C 20	60.8	0.2	85	4	US-08-340-426D-92
C 21	60.8	0.2	85	4	US-08-450-673C-92
C 22	60.8	0.2	85	6	PCT-US95-17111A-92
C 23	60.4	0.2	78	3	US-08-454-557C-70
C 24	60.4	0.2	78	4	US-08-340-426D-70
C 25	60.4	0.2	78	4	US-08-450-673C-70
C 26	60.4	0.2	78	6	PCT-US95-17111A-70
C 27	58.8	0.2	78	3	US-08-454-557C-70

28	58.8	0.2	78	4	US-08-340-426D-70	Sequence 70, Appl
29	58.8	0.2	78	4	US-08-450-673C-70	Sequence 70, Appl
30	58.8	0.2	78	6	PCT-US95-17111A-70	Sequence 70, Appl
C 31	58	0.2	76	3	US-08-454-557C-69	Sequence 69, Appl
C 32	58	0.2	76	4	US-08-340-426D-69	Sequence 69, Appl
C 33	58	0.2	76	4	US-08-450-673C-69	Sequence 69, Appl
C 34	58	0.2	76	6	PCT-US95-17111A-69	Sequence 69, Appl
C 35	55.2	0.2	60	3	US-08-454-557C-57	Sequence 57, Appl
C 36	55.2	0.2	60	4	US-08-340-426D-57	Sequence 57, Appl
C 37	55.2	0.2	60	4	US-08-450-673C-57	Sequence 57, Appl
C 38	55.2	0.2	60	6	PCT-US95-17111A-57	Sequence 57, Appl
C 39	54.6	0.2	94	5	US-08-750-064-6	Sequence 6, Appl
40	53.6	0.2	60	3	US-08-454-557C-60	Sequence 60, Appl
41	53.6	0.2	60	4	US-08-340-426D-60	Sequence 60, Appl
42	53.6	0.2	60	4	US-08-450-673C-60	Sequence 60, Appl
43	53.6	0.2	60	6	PCT-US95-17111A-60	Sequence 60, Appl
44	51.6	0.2	76	3	US-08-454-557C-69	Sequence 69, Appl
45	51.6	0.2	76	4	US-08-340-426D-69	Sequence 69, Appl

ALIGNMENTS

RESULT 1

US-08-481-658B-65/C
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.3%; Score 80; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. No. 1.8e-09;

Matches 89; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 6147 ATCCGAGCAGCTTTGGAGGTCGAGGCGAGTCATCAGGAGTTCAGAGCAGC 6206
|||||
Db 105 ATCCGAGCAGCTTTGGAGGCGAGGCTGGTGATCACAAGGTCAGAGTTTGAGAGCAGC 46
|||||
Qy 6207 CTGACCAAAATGATAAACCCCTGTCTCTACTAAAAATACAACA 6250
|||||
Db 45 CTGGCCAATATGGTGAACCCCTGTCTCTACTAAAAATGTAATAA 2

RESULT 2

US-08-477-504A-65/c
; Sequence 65, Application US/08477504A
; Patent No. 5972353
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,504A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-477-504A-65

Query Match 0.3%; Score 80; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. No. 1.8e-09;
Matches 89; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 6147 ATCCGAGCAGCTTTGGAGGTCGAGGCGAGTCATCAGGAGTTCAGAGCAGC 6206
|||||
Db 105 ATCCGAGCAGCTTTGGAGGCGAGGCTGGTGATCACAAGGTCAGAGTTTGAGAGCAGC 46
|||||
Qy 6207 CTGACCAAAATGATAAACCCCTGTCTCTACTAAAAATACAACA 6250
|||||
Db 45 CTGGCCAATATGGTGAACCCCTGTCTCTACTAAAAATGTAATAA 2

RESULT 3

US-08-486-756A-65/c
; Sequence 65, Application US/08486756A
; Patent No. 5981711
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,756A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3C
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-486-756A-65

Query Match 0.3%; Score 80; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. No. 1.8e-09;
Matches 89; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 6147 ATCCGAGCAGCTTTGGAGGTCGAGGCGAGTCATCAGGAGTTCAGAGCAGC 6206
|||||
Db 105 ATCCGAGCAGCTTTGGAGGCGAGGCTGGTGATCACAAGGTCAGAGTTTGAGAGCAGC 46
|||||
Qy 6207 CTGACCAAAATGATAAACCCCTGTCTCTACTAAAAATACAACA 6250
|||||
Db 45 CTGGCCAATATGGTGAACCCCTGTCTCTACTAAAAATGTAATAA 2

RESULT 4

US-08-485-862B-65/c
; Sequence 65, Application US/08485862B
; Patent No. 5989838
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court

CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485.862B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-485-862B-65

Query Match 0.3%; Score 80; DB 4; Length 105;
Best Local Similarity 85.6%; Pred. No. 1.8e-09;
Matches 89; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 6147 ATCCGAGCACTTTGGAGGTCGAGGCTGATCAGGAGTTCAGAGTTCAGAGCAGC 6206
Db 105 ATCCGAGCACTTTGGAGGTCGAGGCTGATCAGGAGTTCAGAGTTCAGAGCAGC 46

QY 6207 CTGACCAAAATGATGAACCCCTGCTCTACTAAATAACAACA 6250
Db 45 CTGGCCAATATGGTGAACCCCTGCTCTACTAAAGATGTAAAAA 2

RESULT 5
US-08-787-739-65/C
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/787.739
FILING DATE: 24-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/486,756
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/481,658
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,862
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,863
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,077
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-981-0332
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-787-739-65

Query Match 0.3%; Score 80; DB 5; Length 105;
Best Local Similarity 85.6%; Pred. No. 1.8e-09;
Matches 89; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 6147 ATCCGAGCACTTTGGAGGTCGAGGCTGATCAGGAGTTCAGAGTTCAGAGCAGC 6206
Db 105 ATCCGAGCACTTTGGAGGTCGAGGCTGATCAGGAGTTCAGAGTTCAGAGCAGC 46

QY 6207 CTGACCAAAATGATGAACCCCTGCTCTACTAAATAACAACA 6250
Db 45 CTGGCCAATATGGTGAACCCCTGCTCTACTAAAGATGTAAAAA 2

RESULT 6
US-08-481-658B-65
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:

MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-486-756A-65

Query Match 0.3%; Score 79.4; DB 4; Length 105;
Best Local Similarity 84.8%; Pred. No. 2.5e-09;
Matches 89; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 10891 TTTTGTATTTTATTAGAGATGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTC 10950
DB 1 TTTTGTATTTTATTAGAGATGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTC 60
QY 10951 CTGACCTCTGTATCCACCGCTTTGGCCCTCCCAAAAGTCTGGGAT 10995
DB 61 CTGACCTCTGTATCCACCGCTTCGCCCTCCCAAAAGTCTGGGAT 105

RESULT 9
US-08-485-862B-65
; Sequence 65, Application US/08485862B
; Patent No. 5989838
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485.862B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-485-862B-65

Query Match 0.3%; Score 79.4; DB 4; Length 105;
Best Local Similarity 84.8%; Pred. No. 2.5e-09;
Matches 89; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 10891 TTTTGTATTTTATTAGAGATGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTC 10950
DB 1 TTTTGTATTTTATTAGAGATGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTC 60
QY 10951 CTGACCTCTGTATCCACCGCTTTGGCCCTCCCAAAAGTCTGGGAT 10995
DB 61 CTGACCTCTGTATCCACCGCTTCGCCCTCCCAAAAGTCTGGGAT 105
RESULT 10
US-08-787-739-65
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-787-739-65

```
Query Match      0.3%; Score 79.4; DB 5; Length 105;
Best Local Similarity 84.8%; Pred. No. 2.5e-09;
Matches 89; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 10891 TTTTGTGATTTTATTAGAGATGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTC 10950
        ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TTTTGTACATCTTTAGTAGAGACAGGTTTCACCATATTTGCCAGGCTGCTCTCAAACTC 60

Qy 10951 CTGACCTGTCATCCACCGCTTTGGCCTCCCAAGTCTGGGAT 10995
        ||||| ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 CTGACCTGTGATCACACAGCCTCGGCTCCCAAGTCTGGGAT 105

RESULT 11
US-08-454-557C-91
; Sequence 91, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2540
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; US-08-454-557C-91

Query Match      0.2%; Score 65.4; DB 3; Length 84;
Best Local Similarity 86.7%; Pred. No. 3.5e-06;
Matches 72; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 10923 CCATGTTAGCCAGGATGTCGATCTCCTGACCTCGTGATCCACCGCTTTGGCCTCCC 10982
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CCATGTTTCATCAGGCTGGTGTGCGAATCCTGACCTCGTGATCCGCGCCCTCAGCCTCCC 60

Qy 10983 AAAGTCTGGGATTACAGCGCTG 11005
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 AAAGTCTGGGATTACAGCGTG 83

RESULT 12
US-08-340-426D-91
; Sequence 91, Application US/08340426D
; Patent No. 5948634
```

```
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2540
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; US-08-340-426D-91

Query Match      0.2%; Score 65.4; DB 4; Length 84;
Best Local Similarity 86.7%; Pred. No. 3.5e-06;
Matches 72; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 10923 CCATGTTAGCCAGGATGTCGATCTCCTGACCTCGTGATCCACCGCTTTGGCCTCCC 10982
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CCATGTTTCATCAGGCTGGTGTGCGAATCCTGACCTCGTGATCCGCGCCCTCAGCCTCCC 60

Qy 10983 AAAGTCTGGGATTACAGCGCTG 11005
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 AAAGTCTGGGATTACAGCGTG 83

RESULT 13
US-08-450-673C-91
; Sequence 91, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```



```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450.673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-450-673C-91

Query Match 0.2%; Score 65.4; DB 4; Length 84;
Best Local Similarity 86.7%; Pred. No. 3.5e-06;
Matches 72; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 10923 CCATGTTAGCCAGGATGCTCGATCTCCTGACCTCGTGATCCACCGCTTTGGCCTCCC 10982
Db 1 CCATGTTTCATCAGGCTGCTCGAACTCCTGACCTCGTGATCCCGCGCTCAGCCTCCC 60

Qy 10983 AAAGTGCTGGGATTACAGCGGTG 11005
Db 61 AAAGTGCTGGGATTACAAAGCGTG 83

RESULT 14
PCT-US95-17111A-91
; Sequence 91, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: PCT/US95/17111A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-450-673C-91

Query Match 0.2%; Score 65.4; DB 4; Length 84;
Best Local Similarity 86.7%; Pred. No. 3.5e-06;
Matches 72; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 10923 CCATGTTAGCCAGGATGCTCGATCTCCTGACCTCGTGATCCACCGCTTTGGCCTCCC 10982
Db 1 CCATGTTTCATCAGGCTGCTCGAACTCCTGACCTCGTGATCCCGCGCTCAGCCTCCC 60

Qy 10983 AAAGTGCTGGGATTACAGCGGTG 11005
Db 61 AAAGTGCTGGGATTACAAAGCGTG 83

RESULT 14
PCT-US95-17111A-91
; Sequence 91, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: PCT/US95/17111A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-450-673C-91
```

```
;
; TOPOLOGY: both
;
PCT-US95-17111A-91

Query Match 0.2%; Score 65.4; DB 6; Length 84;
Best Local Similarity 86.7%; Pred. No. 3.5e-06;
Matches 72; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 10923 CCATGTTAGCCAGGATGCTCGATCTCCTGACCTCGTGATCCACCGCTTTGGCCTCCC 10982
Db 1 CCATGTTTCATCAGGCTGCTCGAACTCCTGACCTCGTGATCCCGCGCTCAGCCTCCC 60

Qy 10983 AAAGTGCTGGGATTACAGCGGTG 11005
Db 61 AAAGTGCTGGGATTACAAAGCGTG 83

RESULT 15
US-08-454-557C-91/c
; Sequence 91, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-454-557C-91

Query Match 0.2%; Score 62.8; DB 3; Length 84;
Best Local Similarity 85.4%; Pred. No. 1.4e-05;
Matches 70; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 6137 CATGACTGTAAATCCAGCACCTTTGGGAGGTCGAGGAGGCTGATCAGGAGTCAGAGTT 6196
Db 83 CACGCTTGTAATCCAGCACCTTTGGGAGGTCGAGGAGGTCGAGGAGTCAGAGTT 24

Qy 6197 CAAGACACGCTGACCAAAATG 6218
Db 23 CGACACACGCTGTGATGACATG 2
```

Search completed: June 15, 2000~ 04:56:24
Job time: 57714 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2000, 04:50:04 ; Search time 17971.5 Seconds
(without alignments)
-1569.819 Million cell updates/sec

Title: US-08-852-495c-1_copy_56000_85000
Perfect score: 29001
Sequence: 1 ATGACAAAGGCTAGCTGAT.....CAGGAGACTAGAGTTTATT 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_bal.*
- 2: gb_ba2.*
- 3: gb_ba1.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl1.*
- 9: gb_pl2.*
- 10: gb_pr1.*
- 11: gb_pr2.*
- 12: gb_pr3.*
- 13: gb_ro.*
- 14: gb_sy.*
- 15: gb_sy.*
- 16: gb_vl.*
- 17: em_fun.*
- 18: em_hum1.*
- 19: em_hum2.*
- 20: em_in.*
- 21: em_om.*
- 22: em_or.*
- 23: em_ov.*
- 24: em_pat.*
- 25: em_ph.*
- 26: em_pl.*
- 27: em_ro.*
- 28: em_sts.*
- 29: em_sy.*
- 30: em_un.*
- 31: em_vl.*
- 32: gb_htg1.*
- 33: gb_htg2.*
- 34: gb_in1.*
- 35: gb_in2.*
- 36: em_ba1.*
- 37: em_ba2.*
- 38: em_hum3.*
- 39: em_hum4.*
- 40: gb_pr4.*
- 41: gb_htg3.*
- 42: gb_htg4.*
- 43: gb_htg5.*
- 44: gb_htg6.*

- 45: gb_htg7.*
- 46: em_htg1.*
- 47: em_htg2.*
- 48: em_htg3.*
- 49: em_hum5.*
- 50: gb_pl3.*
- 51: gb_pr5.*
- 52: gb_htg8.*
- 53: gb_htg9.*
- 54: gb_htg10.*
- 55: gb_htg11.*
- 56: gb_htg12.*
- 57: gb_htg13.*
- 58: gb_htg14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	92	0.3	108	10	HSLDLRN2
2	92	0.3	108	10	HSLDLRN2
3	88.6	0.3	107	9	HUMALCE162
4	88	0.3	108	11	HSU67803
5	84	0.3	103	9	HUMALCE221
6	83.6	0.3	103	9	HUMALCE221
7	83.8	0.3	104	9	HUMALCE272
8	83	0.3	108	10	HSLDLRD1
9	83	0.3	108	10	HSLDLRD2
10	81.4	0.3	108	10	HSLDLRD1
11	81.4	0.3	108	10	HSLDLRD2
12	78.2	0.3	91	13	HUMUT8164A
13	78.4	0.3	108	10	HSLDLI12
14	77.8	0.3	108	11	HSU67803
15	77.8	0.3	110	9	HUMALCE43
16	76.2	0.3	108	11	HSU67808
17	76.4	0.3	108	11	HSU67808
18	76.4	0.3	110	9	HUMALCE43
19	76	0.3	104	9	HUMALCE272
20	75.8	0.3	107	9	HUMALCE162
21	75.4	0.3	106	13	G32743
22	75.2	0.3	108	11	HSU67804
23	74.4	0.3	103	13	HS8IC8R
24	73.4	0.3	103	13	HS8IC8R
25	72	0.2	90	9	HUMLDLRFL
26	72	0.2	108	10	HSLDLI12
27	71.4	0.2	107	11	HSU67806
28	71.4	0.2	108	11	HSU67804
29	70.8	0.2	91	13	HUMUT8164A
30	70	0.2	107	11	HSU67806
31	70	0.2	108	9	HUMDI03M5
32	70.2	0.2	110	11	HSU67807
33	70.2	0.2	110	11	HSU67807
34	69.4	0.2	97	9	HUMLDLRA2
35	69.4	0.2	100	9	HUMGALNSA
36	68.8	0.2	80	9	HUMBRKFAE
37	68.4	0.2	95	13	HUMUT8002B
38	68.4	0.2	102	13	G32906
39	67.6	0.2	101	10	S79560
40	67.8	0.2	108	9	HUMDI03M5
41	67.2	0.2	84	5	AR051521
42	67.2	0.2	90	9	HUMLDLRFL
43	66.8	0.2	102	13	G32906
44	66	0.2	100	10	HSLAS27
45	66	0.2	100	13	HUMUT931A

ALIGNMENTS

```

RESULT 1
HSLDLRN2
LOCUS      HSLDLRN2      108 bp      DNA      PRI      20-MAY-1992
DEFINITION Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION X05250
VERSION    X05250.1 GI:34337
KEYWORDS   Alu repetitive sequence; low density lipoprotein receptor.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 108)
AUTHORS    Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
TITLE      Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
JOURNAL    Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE    87161901
COMMENT    See X05252 for deletion junction
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES   Location/Qualifiers
            source      1..108
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
            intron      1..108
                        /note="intron XIV fragment"
BASE COUNT 28 a 23 c 39 g 18 t
ORIGIN

Query Match      0.3%; Score 92; DB 10; Length 108;
Best Local Similarity 90.7%; Pred. No. 2.le-06;
Matches 98; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 10989 ACRAAGTGTGGGGTGGTGCACATCCCTGTAGTCCAGCTACTGGGGAGGCTGAG 11048
         1 ACAAATTTAGCCAGCGGTGGTGGCAGGTGCTGTATCCAGCTACTCGGAGGCTGAG 60
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 11049 GCAGGAGAAATGCTGAACTCGGGAGCGGAGGTGTCAGTGCAGCCGAG 11096
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GCAGGAGAAATGCTTGAACCCAGGAGGAGGTTGCAGTGCAGCCGAG 108

RESULT 2
HSLDLRN2/c
LOCUS      HSLDLRN2      108 bp      DNA      PRI      20-MAY-1992
DEFINITION Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION X05250
VERSION    X05250.1 GI:34337
KEYWORDS   Alu repetitive sequence; low density lipoprotein receptor.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 108)
AUTHORS    Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
TITLE      Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
JOURNAL    Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE    87161901
COMMENT    See X05252 for deletion junction
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES   Location/Qualifiers
            source      1..108
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
            intron      1..108
                        /note="intron XIV fragment"
BASE COUNT 28 a 23 c 39 g 18 t
ORIGIN

```

ORIGIN

```

Query Match      0.3%; Score 92; DB 10; Length 108;
Best Local Similarity 90.7%; Pred. No. 2.le-06;
Matches 98; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 24355 CTCGGCTCACTGCAACCTCGGCTCAGCGGTTCAAGCGATTCTCCTGCCCTCCCG 24414
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 108 CTCGGCTCACTGCAACCTCGGCTCCTCGGTTCAAGCAATCTCTCGCTCAGCCTCCG 49
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 24415 AGTAGCTGAGATTACAGGGCGCTGCCACCATCGCCGGCTAAATTTTGT 24462
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 48 AGTAGCTGGGATTACAGGCACCTGCCACCGCTGGCTAATTTTGT 1

RESULT 3
HUMALCE162/c
LOCUS      HUMALCE162      107 bp      ss-RNA      PRI      15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE162.
ACCESSION M87924
VERSION    M87924.1 GI:174871
KEYWORDS   Alu repeat.
SOURCE     Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 107)
AUTHORS    Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE      Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
JOURNAL    J. Mol. Biol. (1992) In press
FEATURES   Location/Qualifiers
            source      1..107
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /cell_line="Ntera2D1"
                        /dev_stage="embryo"
                        /sex="male"
                        /tissue_type="carcinoma"
BASE COUNT 28 a 30 c 35 g 14 t
ORIGIN

Query Match      0.3%; Score 88.6; DB 9; Length 107;
Best Local Similarity 91.3%; Pred. No. 8e-06;
Matches 94; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 11447 TGTGTGAGACGGAGTCTTGTCTGTGTGCCAGGCTGAGTGCAGTGGTGTGATCTCCG 11506
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 107 TTTTGTGAGACGGAGTCTCGCTCTGTGCGCCAGGCTGAGTGCAGTGGCGGATCTCGG 48
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 11507 TCAGTGAAGCTCCGCTCCCGGATTACGCCATTCTCTGCC 11549
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 47 TCAGTGAAGCTCCGCTCCCGGTTTCACGCCATTCTCTGCC 5

RESULT 4
HSU67803/c
LOCUS      HSU67803      108 bp      RNA      PRI      01-AUG-1997
DEFINITION Human small cytoplasmic Alu transcript.
ACCESSION U67803
VERSION    U67803.1 GI:2289917
KEYWORDS   Alu.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 108)
AUTHORS    Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE      cDNAs derived from primary and small cytoplasmic Alu (scAlu)
            transcripts
JOURNAL    J. Mol. Biol. 271 (2), 222-234 (1997)

```

```

MEDLINE 97415756
REFERENCE 2 (bases 1 to 108)
AUTHORS Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The Children's Hospital of Philadelphia, 1004F Abramson Research Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES
    Location/Qualifiers
    source 1..108
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="TscAlu2"
    repeat_region 1..108
        /note="scAlu"
        /rpt_family="Alu"
        /rpt_type="dispersed"
BASE COUNT 23 a 39 c 30 g 16 t
ORIGIN
Query Match 0.3%; Score 88; DB 11; Length 108;
Best Local Similarity 94.8%; Pred. No. 1e-05;
Matches 91; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 11616 GTAGATACGGGTTTACCTTTGTTAAACCAAGGATGCTCGATCTCTGACCTCGTGATCG 11675
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 97 GTAGAGACGGGTTTACCTTTGTTAAACCAAGGATGCTCGATCTCTGACCTCGTGATCC 38
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 11676 GCCCGCTCAGCTCCCAAGTGTGGGATTACAGG 11711
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 37 GCCCGCTCGGCTCCCAAGTGTGGGATTACAGG 2
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 5
HUMALCE221
LOCUS HUMALCE221 103 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
    Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
    Location/Qualifiers
    source 1..103
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /cell_line="Ntera2D1"
        /dev_stage="embryo"
        /sex="male"
        /tissue_type="carcinoma"
BASE COUNT 25 a 27 c 33 g 18 t
ORIGIN
Query Match 0.3%; Score 84; DB 9; Length 103;
Best Local Similarity 90.0%; Pred. No. 4.9e-05;
Matches 90; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 11017 TGCCTGTAGTCCAGCTACTGGGAGGCTGAGCAGAGAGTGTGTAACCTGGGAGGC 11076
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4 TGCCTGTATCCAGCTACACGGAGCTTAAGCAGGAGAGTGTGTAACCTGGGAGGC 63
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 11077 GGAGGTTCAGTGCAGCCGAGATGCGCCACTGCATCCAG 11116
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 64 GGAGGTTCAGTGCAGCCGAGATGCTGCCATTGCACTCCAG 103
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

```

RESULT 6
HUMALCE221/c
LOCUS HUMALCE221 103 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
    Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
    Location/Qualifiers
    source 1..103
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /cell_line="Ntera2D1"
        /dev_stage="embryo"
        /sex="male"
        /tissue_type="carcinoma"
BASE COUNT 25 a 27 c 33 g 18 t
ORIGIN
Query Match 0.3%; Score 83.6; DB 9; Length 103;
Best Local Similarity 90.8%; Pred. No. 5.8e-05;
Matches 89; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 24335 CTGGAGTGTGTCGGCAGCTTCGGCTCACTGCAACCTCCGCTCAGGGTTCAAGCGAT 24394
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 103 CTGGAGTGTCAATGGCAGCATCTGGCTCACTGCAACCTCCGCTCAGGGTTCAAGCGAT 44
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 24395 TCTCTGCTCCCTCCGCTCCGAGTAGCTGAGATTACAGG 24432
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 43 TCTCTGCTCCCTAGCTCCGCTGCTAGCTGGGATTACAGG 6
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 7
HUMALCE272/c
LOCUS HUMALCE272 104 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE272.
ACCESSION M87895
VERSION M87895.1 GI:174875
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
    Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 104)
AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
    Location/Qualifiers
    source 1..104
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /cell_line="Ntera2D1"
        /dev_stage="embryo"
        /sex="male"
        /tissue_type="carcinoma"
BASE COUNT 22 a 26 c 37 g 19 t
ORIGIN
Query Match 0.3%; Score 83.8; DB 9; Length 104;
Best Local Similarity 88.3%; Pred. No. 5.3e-05;
Matches 91; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```



```
source 1. .108
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature complement(<1..65)
/notes="Alu repeat"
intron 1. .108
/notes="intron XII fragment"
BASE COUNT 21 a 38 c 20 g 29 t
ORIGIN

Query Match 0.3%; Score 78.4; DB 10; Length 108;
Best Local Similarity 84.6%; Pred. No. 0.00045;
Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 10992 AAAGTACCTGGCGGTGGGCACATGCCTTACTCCAGCTACTGGGAGGCTGAGGCA 11051
|||||
Db 105 AAAATCAGCGCGGTGGGCACATGCTTGTATCCAGCTACTAAGGAGGCTGAGGCA 46
|||||

Qy 11052 GGAGAAATGCTTGAACTCGGGAGGCGGAGGTTCGAGTGAGCGCA 11095
|||||
Db 45 GGAAATGTTTGAACCCAGGAGGCGAGAGGTGTGTGAGGCGCA 2
|||||

RESULT 14
HSU67803 108 bp RNA PRI 01-AUG-1997
LOCUS
DEFINITION Human small cytoplasmic Alu transcript.
ACCESSION U67803
VERSION U67803.1 GI:2289917
KEYWORDS Alu.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE cDNAs derived from primary and small cytoplasmic Alu (scAlu)
transcripts
JOURNAL J. Mol. Biol. 271 (2), 222-234 (1997)
MEDLINE 97415756
REFERENCE 2 (bases 1 to 108)
AUTHORS Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
Children's Hospital of Philadelphia, 1004F Abramson Research
Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES
source
1. .108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TscAlu2"
repeat_region 1. .108
/notes="scAlu"
/rpt_family="Alu"
/rpt_type="dispersed"
BASE COUNT 23 a 39 c 30 g 16 t
ORIGIN

Query Match 0.3%; Score 77.8; DB 11; Length 108;
Best Local Similarity 87.6%; Pred. No. 0.00057;
Matches 85; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 25011 GCCTATATCCAGCTCTTTGGGAGGCTTAGCGGGTGGATCAGCAGGTTCAGGAGTTCAA 25070
|||||
Db 1 GCCTGTATCCAGCACTTTTGGGAGCGCGAGCGGGCGGATCAGCAGGTTCAGGAGATCGA 60
|||||

Qy 25071 GACCGCTCGCCCAAGATGGTGAATCCCGTCTCTAC 25107
|||||
Db 61 GACCATCTGCGCTAACACAGGTGAACCCCGTCTCTAC 97
|||||
```

```
RESULT 15
HUMALCE43 110 bp ss-RNA PRI 15-APR-1994
LOCUS
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE43.
ACCESSION M87900
VERSION M87900.1 GI:174876
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 110)
AUTHORS Sinnott,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
source
1. .110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Ntera2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT 27 a 31 c 34 g 18 t
ORIGIN

Query Match 0.3%; Score 77.8; DB 9; Length 110;
Best Local Similarity 83.8%; Pred. No. 0.00057;
Matches 88; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 24991 GGCTGGGCGGCTGCCTCACGCTATATCCAGCTCTTTGGGAGGCTAGCGGGTGA 25050
|||||
Db 5 GGCGGGCGCGGTGGCTCACACCTACATCCCGGACCTTTGGGAGGCCAAGCAGGAGGA 64
|||||

Qy 25051 TCACGAGCTCAGGAGTTCAAGACCAGCAGCTCGCCAGATGGTGA 25095
|||||
Db 65 TTGCGAGCTCAGTAGTTCAGAGACCAGCAGCTGCTAACATGACGAAA 109
|||||

Search completed: June 15, 2000, 13:12:08
Job time: 88412 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: June 15, 2000, 05:38:04 ; Search time 593.79 seconds
(without alignments)
12219.506 Million cell updates/sec

Title: US-08-852-495C-1_COPY_56000_85000
Perfect score: 29001
Sequence: 1 ATGAACAAGGCTGACTGAT.....CAGGAGACTAGACTTTTATT 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 433070

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	71	0.2	100	1 T24892	Human gene signatu
C 2	70.4	0.2	100	1 T24892	Human gene signatu
C 3	68.8	0.2	108	1 T26828	Human gene signatu
C 4	66.8	0.2	108	1 X12095	Human biallelic po
C 5	64.8	0.2	108	1 X12095	Human biallelic po
C 6	64.2	0.2	91	1 T25854	Human gene signatu
C 7	64.4	0.2	108	1 T26828	Human gene signatu
C 8	63.6	0.2	103	1 T26213	Human gene signatu
C 9	62.8	0.2	103	1 T20927	Human gene signatu
C 10	62.4	0.2	103	1 T20927	Human gene signatu
C 11	61	0.2	91	1 T25854	Human gene signatu
C 12	61	0.2	103	1 T26213	Human gene signatu
C 13	59.6	0.2	100	1 X12087	Human biallelic po
C 14	59.6	0.2	100	1 X12085	Human biallelic po
C 15	59.6	0.2	100	1 X12086	Human biallelic po
C 16	59	0.2	93	1 T22572	Human gene signatu
C 17	59.2	0.2	108	1 T25009	Human gene signatu
C 18	58.4	0.2	110	1 T26728	Human gene signatu
C 19	57.4	0.2	97	1 T26288	Human gene signatu
C 20	57.4	0.2	110	1 T26288	Human gene signatu
C 21	57	0.2	100	1 X12087	Human biallelic po
C 22	57	0.2	100	1 X12085	Human biallelic po
C 23	57	0.2	100	1 X12086	Human biallelic po
C 24	57	0.2	108	1 T25009	Human gene signatu
C 25	56	0.2	93	1 T25688	Human gene signatu
C 26	55.8	0.2	95	1 T23131	Human gene signatu
C 27	55	0.2	109	1 T23895	Human gene signatu
C 28	54.4	0.2	109	1 T23895	Human gene signatu
C 29	53.6	0.2	99	1 T20931	Human gene signatu
C 30	53.2	0.2	75	1 T22841	Human gene signatu
C 31	53.2	0.2	93	1 T22572	Human gene signatu
C 32	52.6	0.2	70	1 N60231	Normal chromosome
C 33	52.2	0.2	97	1 T26728	Human gene signatu
C 34	51.4	0.2	53	1 Q33621	Microsatellite seq

C 35	51.6	0.2	75	1 T22841	Human gene signatu
C 36	51.4	0.2	93	1 T24259	Human gene signatu
C 37	51	0.2	93	1 T24259	Human gene signatu
C 38	50.4	0.2	85	1 T26182	Human gene signatu
C 39	50.4	0.2	106	1 Q95210	Simple tandem repe
C 40	50	0.2	87	1 T21566	Human gene signatu
C 41	49.6	0.2	81	1 T24093	Human gene signatu
C 42	49.6	0.2	84	1 T25848	Human gene signatu
C 43	49.6	0.2	88	1 V39744	Microsatellite ana
C 44	49.6	0.2	93	1 T25688	Human gene signatu
C 45	49.8	0.2	100	1 T25604	Human gene signatu

ALIGNMENTS

RESULT 1
T24892/C
ID T24892 standard; cDNA to mRNA; 100 BP.
AC T24892;
DT 05-NOV-1996 (first entry)
DE Human gene signature HUMGS06998.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W0951472-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1720; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;

Query Match 0.2%; Score 71; DB 1; Length 100;
Best Local Similarity 81.6%; Pred. No. 0.015;
Matches 80; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 17609 TTTTGTGTTTGAATAGAGTCTCGCTCTCTACCCAGCTGGAGTGCAGTGGCGCAATCT 17668

Db 99 TTGTTGTTTGAATAGAGTCTCGCTCTCTACCCAGCTGGAGTGCAGTGGCGCAATCT 40

QY 17669 CAGCTCACTGCAACGTCGCCCTCTCGGGTTCAAGTAT 17706

Db 39 CAGCTNATTCGAATCTCGCTCTCCAGGTTCAAGCGAT 2

RESULT 2

T24892
ID T24892 standard; cDNA to mRNA; 100 BP.


```
SQ Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;

Query Match 0.2%; Score 66.8; DB 1; Length 108;
Best Local Similarity 81.5%; Pred. No. 0.063;
Matches 88; Conservative 1; Mismatches 18; Indels 1; Gaps 1;

Oy 7604 CTGTAATCCAGCAC-TTTGGAGGCTGAGTGGATGATCAGCTGAGTGGAGTTTG 7662
II IIIIIIIIIII IIIIIIIII III IIIIIII III IIIIIII III IIIIIII
Db 108 CTATATCCAGCACTTTTGGAGGCCAAGCAGACGGATCACTTGAAGTCAAGGATCG 49
II IIIIIIIIIII IIIIIIIII III IIIIIII III IIIIIII III IIIIIII

Oy 7663 AGACCAAGCTGCCCAACATGTTAAACCCCATCTGTTACTTAAATAACA 7710
IIIIII IIIIIIIIIII IIIIIII IIIIIII IIIIIII IIIIIII IIIIIII
Db 48 AGACCATCTGCCCAACAYAGGAAACCTCATCTCTACAAAAAGACA 1

RESULT 5
X12095
ID X12095 standard; DNA; 108 BP.
AC X12095;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment TIGR-A003M18a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medication;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-TUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
CC Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;

Query Match 0.2%; Score 64.8; DB 1; Length 108;
Best Local Similarity 81.1%; Pred. No. 0.12;
Matches 86; Conservative 1; Mismatches 18; Indels 1; Gaps 1;

Oy 19552 TTTTATAGATAGAGTGGGTTTACCATGCTGGCCAGGCTGGTTCGAACCTCCTGACT 19611
IIIIII IIIIIIIII IIIII IIIIIII IIIIIII IIIIIII IIIIIII IIIIIII
Db 3 TCTTTTGTAGATAGAGTGGTTCCTTGTGGCCAGGATGCTCTCGAATCCTCCTGACT 62
IIIIII IIIIIIIII IIIII IIIIIII IIIIIII IIIIIII IIIIIII IIIIIII

Oy 19612 CAGGCGATCTCCCGCCCTCAGCCCTCCCAA-CTGCTAGGATTACAG 19656
II IIIII IIIII IIIII IIIIIII IIIIIII IIIIIII IIIIIII IIIIIII
Db 63 CAAGTGATCCGCTCTGCCTTGGCCCTCCCAAAAGTGCTGGGATTATAG 108
IIIIII IIIIIIIII IIIII IIIIIII IIIIIII IIIIIII IIIIIII IIIIIII
```

```
RESULT 6
T25854
ID T25854 standard; cDNA to mRNA; 91 BP.
AC T25854;
DT 22-OCT-1996 (first entry)
DE Human gene signature HUMGS08084.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1944; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
CC Sequence 91 BP; 18 A; 22 C; 28 G; 18 T;

Query Match 0.2%; Score 64.2; DB 1; Length 91;
Best Local Similarity 80.9%; Pred. No. 0.15;
Matches 72; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Oy 4120 ATCACTTAACTCAGGAGGCGAGGTTGTCAGTGCAGTGCATCACCATTGCCTCCAG 4179
IIIIII IIIIIII IIIIIII IIIII IIIIIII IIIIIII IIIIIII IIIIIII
Db 2 ATCACTTGAGCCTTAGGAGGCGAGNGGTTCAAGTGCAGTGCAGTGCCTCCTCGCTCCAG 61
IIIIII IIIIIII IIIIIII IIIII IIIIIII IIIIIII IIIIIII IIIIIII

Oy 4180 CTTGGGTGCACAGTGTGAGACTCTGTCTCA 4208
III IIIIIII IIIII IIIIIII IIIIIII IIIIIII IIIIIII IIIIIII
Db 62 CTNGGTGCACGCGTGAGANNCTGTCTCA 90
IIIIII IIIIIII IIIII IIIIIII IIIIIII IIIIIII IIIIIII IIIIIII

RESULT 7
T26828
ID T26828 standard; cDNA to mRNA; 108 BP.
AC T26828;
DT 14-NOV-1996 (first entry)
DE Human gene signature HUMGS09078.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
```

PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 2182; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 18 A; 33 C; 23 G; 28 T;

Query Match 0.2%; Score 64.4; DB 1; Length 108;
Best Local Similarity 80.4%; Pred. No. 0.14;
Matches 74; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 11655 GATCTCTGACCTGATCGCCGCGCTCAGCTCCCAAGTCTGGATTACAGGAGT 11714
|||||
Db 1 GATCTCTGACCTGATCGCCGCGCTCAGCTCCCAAGTCTGGATTACAGGAGT 60

QY 11715 GAGCCACTGCGCCGCGCGCTTTTTTTTTT 11746
|||||
Db 61 GAGCCACTGCGCGCGCTTTTTTTTATTCTTAT 92

RESULT 8

T26213 ID T26213 standard; cDNA to mRNA; 103 BP.
AC T26213;
DT 13-NOV-1996 (first entry)
DE Human gene signature HUMGS08452.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K. Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 2029; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

SQ Sequence 103 BP; 33 A; 21 C; 25 G; 23 T;

Query Match 0.2%; Score 63.6; DB 1; Length 103;
Best Local Similarity 76.5%; Pred. No. 0.19;
Matches 78; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 10154 GATCTCTTGAGCTAGAAAGTTTGGGAGCAGTGAGCTATGATTATCCCACTGCACATCCA 10213
|||||
Db 1 GATCACTTGAGTCCAGGAGTTTGGTTACAGTGGATGATGATGGCCACCACTGCATCCA 60

QY 10214 GCCTGGCCAAATGCAAAATCTCTCAAAACAAAAACA 10255
|||||
Db 61 GCCTGGCCCAAGTAAGAACAATGCTCTTAAAGAAAAAAA 102

RESULT 9

T20927 ID T20927 standard; cDNA to mRNA; 103 BP.
AC T20927;
DT 24-JUL-1996 (first entry)
DE Human gene signature HUMGS02180.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K. Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 758-759; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 103 BP; 22 A; 27 C; 21 G; 31 T;

Query Match 0.2%; Score 62.8; DB 1; Length 103;
Best Local Similarity 76.0%; Pred. No. 0.24;
Matches 76; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 17704 GATTCCTCCCTCAGCTCCCAAGTAGCTGGGATTACAGTCACGCCACCATGCCGAG 17763
|||||
Db 1 GATCCTCCCACTCCCACTCCCAAGTAGCTGGCTACAGTGTGTGCCCATGTCCAG 60

QY 17764 CTAATTTTGTATTTTGAAGACAGCGAATTCACCATG 17803
|||||
Db 61 CTGATTTTNGTATTTTNAAGTAGGACACAGTATTTCTCCATG 100

RESULT 10

T20927/c ID T20927 standard; cDNA to mRNA; 103 BP.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2000, 23:50:22 ; Search time 8476.34 Seconds
(without alignments)
13867.722 Million cell updates/sec

Title: US-08-852-495C-1_COPY_56000_85000
Perfect score: 29001
Sequence: 1 ATGACAAAGGCTCACTGAT.....CAGGAGACTAGACTTTTATT 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : EST:
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result % Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	94.8	0.3	106	37	AA703692	ag81a10.r
2	91.4	0.3	109	30	AA243009	zr25h02.s
3	91.4	0.3	109	84	BI7434	345K2.TVB C
4	88.8	0.3	100	42	AI077628	oy26f04.s
5	88.8	0.3	101	39	AA835205	ak64h01.s
6	88.6	0.3	103	84	B48914	RPC111-4A12
7	88.2	0.3	109	84	BI7434	345K2.TVB C
8	88.2	0.3	109	94	AQ028426	CIT-HSP-2
9	88.2	0.3	110	94	AQ003188	RPC111-ID
10	87.6	0.3	110	106	AQ386882	RPC111-13
11	86.8	0.3	106	105	AQ264176	CITBI-EI-
12	86.6	0.3	110	30	AA244245	nc07a04.s
13	85.8	0.3	107	24	H67040	yu68c01.r1
14	85.8	0.3	108	84	B65160	CIT-HSP-201
15	85.2	0.3	106	38	AA812141	ob48h02.s
16	85.2	0.3	110	39	AA897366	am06h02.s
17	84.6	0.3	107	35	AA565533	nk42b11.s
18	84.6	0.3	103	108	AQ35244	RPCI-11-3
19	84.4	0.3	103	38	AA807640	nx08b05.s
20	84.2	0.3	105	105	AQ282107	RPC111-94
21	84.4	0.3	110	106	AQ386882	RPC111-13
22	83.8	0.3	103	94	AQ028649	CIT-HSP-2
23	83.8	0.3	103	108	AQ35244	RPCI-11-3
24	83.6	0.3	106	63	AI991750	wt48e01.x
25	83.8	0.3	107	33	AA385808	EST99495
26	84	0.3	109	30	AA243009	zr25h02.s
27	83.4	0.3	101	35	AA583697	nn58f10.s
28	83.6	0.3	106	63	AI991750	wt48e01.x
29	83	0.3	91	38	AA780764	ac68f12.s
30	83.2	0.3	105	30	AA218889	zg15d04.s
31	83.2	0.3	107	33	AA385808	EST99495
32	82.8	0.3	103	108	AQ382186	RPCI-11-4
33	82.8	0.3	103	108	AQ382186	RPCI-11-4
34	83	0.3	107	35	AA565533	nk42b11.s
35	82.4	0.3	101	35	AA583697	nn58f10.s
36	82.6	0.3	105	28	AA078003	7H12D08 C
37	82.8	0.3	105	109	AQ637292	RPCI-11-4
38	82.8	0.3	110	109	AQ634950	RPCI-11-4
39	82.2	0.3	103	108	AQ334922	RPCI-11-3
40	82	0.3	106	105	AQ264176	CITBI-EI-
41	82.2	0.3	106	105	AQ282340	RPC111-80
42	82	0.3	106	108	AQ544957	CITBI-EI-
43	81.8	0.3	109	94	AQ028426	CIT-HSP-2
44	81.8	0.3	109	105	AQ265749	CITBI-EI-
45	81.4	0.3	99	34	AA486800	abl9a06.r

ALIGNMENTS

RESULT 1
AA703692
LOCUS ag81a10.r1 Stratagene hNT neuron (#937233) EST 24-DEC-1997
DEFINITION IMAGE:1140858 5' similar to contains Alu repetitive element; mRNA sequence.
ACCESSION AA703692 106 bp mRNA
VERSION AA703692.1 GI:2713610
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 106)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1397630.

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 53.

FEATURES

source
Location/Qualifiers

1..106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1140858"
/clone_lib="Stratagene hNT neuron (#937233)"
/dev_stage="hNT neurons"
/lab_host="SOLR (kanamycin resistant)"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT.
Differentiated, post mitotic hNT neurons. Average insert size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATCGGCACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' "

BASE COUNT

19 a 29 c 29 g 29 t

Query Match

0.3%; Score 94.8; DB 37; Length 106;
Best Local Similarity 93.4%; Pred. No. 0.044;
Matches 99; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 11610 TTTTCTAGTAGACGGGTTTCACTTGTGTTAACCCAGGATGCTCGATCTCTGACCTCG 11669

Db 1 TTTTCTAGTAGACGGGTTTCACTTGTGTTAACCCAGGATGCTCGATCTCTGACCTCG 60

QY 11670 TGATCGGCCCGCTCAGCTCCCAAGTCGTTGGGATACAGGAGTG 11715

Db 61 TGATCGGCCCGCTCAGCTCCCAAGTCGTTGGGATACAGGAGTG 106

RESULT 2

AA243009 109 bp mRNA EST 11-MAR-1998
LOCUS zr25h02.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
DEFINITION cDNA clone IMAGE:664467 3' similar to contains Alu repetitive element; contains element LTR1 repetitive element; mRNA sequence.
ACCESSION AA243009
VERSION AA243009.1 GI:1873869
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 109)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Dec 3, 1996 this sequence version replaced gi:1136869.

Contact: Wilson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: estewatson.wustl.edu

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 1127 Std Error: 0.00
Seq primer: -4ml3 fwd. ET from Amersham
High quality sequence stop: 102.

FEATURES
source

Location/Qualifiers
1. .109
/organism="Homo sapiens"
/db_xref="GDB:5426481"
/db_xref="taxon:9606"
/clone="IMAGE:664467"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/tissue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="organ: brain; Vector: plasmid; SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Uninduced, exponentially growing neuroepithelial
cells (Ntera-2/cl.D1). Average insert size: 1.0 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCGCAGG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3"

BASE COUNT 19 a 30 c 30 g 30 t
ORIGIN

Query Match 0.3%; Score 91.4; DB 30; Length 109;
Best Local Similarity 89.9%; Pred. No. 0.11; Mismatches 0; Gaps 0;
Matches 98; Conservative 0; Indels 0; Gaps 0;

QY 11607 GATTTTATAGATACGGGGTTTACCTTTGTTAACGAGATGCTCGATCTCCTGACC 11666
||||| 11607 GATTTTATAGATACGGGGTTTACCTTTGTTAACGAGATGCTCGATCTCCTGACC 11666
Db 1 GATTTTATAGATACGGGGTTTACCTTTGTTAACGAGATGCTCGATCTCCTGACC 60
||||| 11607 GATTTTATAGATACGGGGTTTACCTTTGTTAACGAGATGCTCGATCTCCTGACC 11666
QY 11667 TCGTATCGCGCGCTCAGCTCCCAAGTCTGGGATTACAGGAGTG 11715
||||| 11667 TCGTATCGCGCGCTCAGCTCCCAAGTCTGGGATTACAGGAGTG 11715
Db 61 TCGTATCGCGCGCTCAGCTCCCAAGTCTGGGATTACAGGAGTG 109
||||| 11667 TCGTATCGCGCGCTCAGCTCCCAAGTCTGGGATTACAGGAGTG 109

RESULT 3
B17434/c

LOCUS B17434 109 bp DNA GSS 04-JUN-1998
DEFINITION 345K2.TVB C1978SKA1 Homo sapiens genomic clone A-345K02, genomic
survey sequence.
ACCESSION B17434
VERSION B17434.1 GI:2125183
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 109)
AUTHORS Adams, M.D., Kelley, J.M., Rounsley, S.R. and Venter, J.C.
TITLE Use of a BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other GSSs: 345K02.TP 345K02.TPB
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: 77
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1. .109
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="A-345K02"
/clone_lib="C1978SKA1"
/sex="Female"
/cell_type="Fibroblast"
/note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
Caltech Human BAC Library A1"

BASE COUNT 24 a 30 c 31 g 24 t

ORIGIN

Query Match 0.3%; Score 91.4; DB 84; Length 109;
Best Local Similarity 89.9%; Pred. No. 0.11; Mismatches 0; Gaps 0;
Matches 98; Conservative 0; Indels 0; Gaps 0;

QY 19558 TAGTAGAGATGGGTTTACCATCTGGCCAGGCTGCTCGAATCTCTGACCTCAGGCG 19617
||||| 19558 TAGTAGAGATGGGTTTACCATCTGGCCAGGCTGCTCGAATCTCTGACCTCAGGCG 19617
Db 109 TAGTTGAGACGGGTTTACCATCTGGCCAGGCTGCTCGAATCTCTGACCTCAGGCG 50
||||| 109 TAGTTGAGACGGGTTTACCATCTGGCCAGGCTGCTCGAATCTCTGACCTCAGGCG 50

QY 19618 ATCTGCCGCCCTCAGCTCCCAAGTCTAGGATTACAGCGCTGAGCCCA 19666
||||| 19618 ATCTGCCGCCCTCAGCTCCCAAGTCTAGGATTACAGCGCTGAGCCCA 19666
Db 49 ATCCGCCCATCAGCTCCCAAGTCTAGGATTATAGTATGATGAGCCCA 1
||||| 49 ATCCGCCCATCAGCTCCCAAGTCTAGGATTATAGTATGATGAGCCCA 1

RESULT 4
A1077628/c
LOCUS A1077628 100 bp mRNA EST 24-SEP-1998
DEFINITION oy26f04.s1 Soares_senescent_fibroblasts_NbHSF Homo sapiens CDNA
clone IMAGE:1666975 3' similar to gb.X57130_cds1 HISTONE H1D
(HUMAN);, mRNA sequence.
ACCESSION A1077628
VERSION A1077628.1 GI:3412036
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 100)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2153443.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 820 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .100
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1666975"
/clone_lib="Soares_senescent_fibroblasts_NbHSF"
/tissue_type="senescent_fibroblast"
/lab_host="BHL08 (ampicillin resistant)"
/note="Vector: pT7T3D (Pharmacia) with a modified
polylinker V_TYPE: phagemid; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5'
TCTTACCAATCTGAAGTCGGACGCCGCGCATTTTTTTTTTTTTTTT 3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo."

BASE COUNT 15 a 29 c 28 g 28 t

Query Match 0.3%; Score 88.8; DB 42; Length 100;
Best Local Similarity 93.0%; Pred. No. 0.22; Mismatches 0; Gaps 0;
Matches 93; Conservative 0; Indels 0; Gaps 0;

QY 22054 AAGAGCCTGGTACGAGGCGCACTCTGGTCAACGAAAGGACCGCTGCTTGGCTCC 22113
||||| 22054 AAGAGCCTGGTACGAGGCGCACTCTGGTCAACGAAAGGACCGCTGCTTGGCTCC 22113
||||| 22054 AAGAGCCTGGTACGAGGCGCACTCTGGTCAACGAAAGGACCGCTGCTTGGCTCC 22113

Db 100 AAGAGCCTGTGGCAAGCGGACCTCTAGTGCACAAAGCAAGCACCCTGCTGTCTGGCTCC 41

Qy 22114 TTTAACTCAACAAGAGCGAGCTCCGGGGAAGCCAGC 22153

Db 40 TTTAACTCAACAAGAGCGGCGCTCCGGGACCCAGC 1

RESULT 5

LOCUS AA835205 101 bp mRNA EST 23-FEB-1998

DEFINITION ak64h01.s1 Barsstead pancreas HPLRB1 Homo sapiens cDNA clone IMAGE:1412689 3' similar to contains Alu repetitive element; contains element KER repetitive element ;, mRNA sequence.

ACCESSION AA835205

VERSION AA835205.1 GI:2908933

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 101)

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

TITLE WashU-NCI human EST Project

JOURNAL Unpublished (1997)

COMMENT On Nov 29, 1993 this sequence version replaced gi:636191.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40m13 fwd. ET from Amersham.

FEATURES

source

1..101

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1412689"

/clone_lib="Barsstead pancreas HPLRB1"

/sex="female"

/dev_stage="adult, 34 years"

/lab_host="DH10B"

/note="Organ: pancreas; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACGAATCTGAAGTGGGAGCGCGCGCTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [AATTCGATCCCTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed by Bob Barsstead."

BASE COUNT 14 a 36 c 27 g 24 t

ORIGIN

Query Match 0.3%; Score 88.8; DB 39; Length 101;

Best Local Similarity 93.0%; Pred. No. 0.22;

Matches 93; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 11453 GAGACGGAGTCTTGCTCTGTGTCGCCAGCTGGAGTGCAGTGTGTGATCTCGCTCACTG 11512

Db 2 GAGACGGAGTCTTCACTGTCTGCCAGCTGGAGTGCAGTGTGTGATCTCGCTCACTG 61

Qy 11513 CAAGCTCCGCTCCCGGATTCACGCCATTCCTCGCTCA 11552

Db 62 CAAGCTCCGCTCCCGGATTCACGCCATTCCTCGCTCA 101

RESULT 6

B48914/c

LOCUS B48914 103 bp DNA GSS 08-APR-1999

DEFINITION RPC111-4A12_TP RPC1-11 Homo sapiens genomic clone RPC1-11-4A12, genomic survey sequence.

ACCESSION B48914

VERSION B48914.1 GI:2601151

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 103)

AUTHORS Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and Venter, J.C.

TITLE Use of BAC End Sequences for Sequence-Ready Map Building

JOURNAL Unpublished (1997)

COMMENT Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are derived from the human BAC library RPC1-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html

Seq primer: SP6

Class: BAC ends.

FEATURES

source

1..103

/organism="Homo sapiens"

/db_xref="GDB:7501163"

/db_xref="taxon:9606"

/clone="RPC1-11-4A12"

/clone_lib="RPC1-11"

/sex="Male"

/cell_type="Lymphocytes"

/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI; RPC111 Human Male BAC Library"

BASE COUNT 30 a 28 c 30 g 15 t

ORIGIN

Query Match 0.3%; Score 88.6; DB 84; Length 103;

Best Local Similarity 91.3%; Pred. No. 0.23;

Matches 94; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 11606 TGTATTTTCTAGATAGCGGGTTTCACTTTGTAAACAGGATGCTCTCGATCTCTGAC 11665

Db 103 TGTATTTTCTAGATAGCGGGTTTCACTTTGTAGCGGATGCTCTCGATCTCTGAC 44

Qy 11666 CTCGTGATCGCGCGCTCGACCTCCCAAGTCTGGGATTAC 11708

Db 43 CTCGTGATCGCGCGCTCGCGCTCCCAAGTCTGGGCTTAC 1

RESULT 7

LOCUS B17434 109 bp DNA GSS 04-JUN-1998

DEFINITION B17434_TV8 C17978SK1 Homo sapiens genomic clone A-345K02, genomic survey sequence.

ACCESSION B17434

VERSION B17434.1 GI:2125183

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 109)

AUTHORS Adams,M.D., Kelley,J.M., Rounsley,S.R. and Venter,J.C.
TITLE Use of a BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other.GSSs: 345K02.TP 345K02.TPB
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.

FEATURES Location/Qualifiers
source 1..109
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="A-345K02"
/clone_lib="CIT978SKA1"
/sex="Female"
/cell_type="Fibroblast"
/note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
Caltech Human BAC Library A1"
BASE COUNT 24 a 30 c 31 g 24 t
ORIGIN

Query Match 0.3%; Score 88.2; DB 84; Length 109;
Best Local Similarity 88.1%; Pred. No. 0.25;
Matches 96; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 7594 TGGCTCATGCTGTAAATCCAGCACTTTGGGAGGCTGAGTGAGTGCACCTGAGGTT 7653
DB 1 TGGCTCATACCTATTAATCTAGCACTTTGGGAGGCTGATGTGGCGGATCACTGAGGTC 60
QY 7654 GGGAGTTTGAGACCACTGGCCCAATGGTAAACCCCATCTCTACTA 7702
DB 61 GGGAGTTTGAGACCACTGGCCCAATGGTAAACCCCGTCTCAACTA 109

RESULT 8
LOCUS A0028426 109 bp DNA GSS 30-JUN-1998
DEFINITION CIT-HSP-2313G15.TF CIT-HSP Homo sapiens genomic clone 2313G15,
genomic survey sequence.
ACCESSION A0028426
VERSION A0028426.1 GI:3268648
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 109)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21

Class: BAC ends.
FEATURES Location/Qualifiers
source 1..109
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2313G15"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT 19 a 36 c 25 g 29 t
ORIGIN

Query Match 0.3%; Score 88.2; DB 94; Length 109;
Best Local Similarity 88.1%; Pred. No. 0.25;
Matches 96; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 17610 TTTTGTGAAATAGAGTCTGCTGTCAACCCAGGCTGGAGTGCAGTGGCGCAATCTC 17659
DB 1 TTGTTTCTGAGCGGACTCTCAGTCTGTCAACCCAGGCTGGAGTGCAGTGGCGCAGTCTG 60
QY 17670 AGCTCACTGCAACGTCGCCCTCTCGGTTCAAGTGAATTCCTCGCTCA 17718
DB 61 AGCTCACTGCAACCTCCACCTCTCGGTTCAAGCGATTCCTCGCTCA 109

RESULT 9
LOCUS A0003188 110 bp DNA GSS 14-APR-1999
DEFINITION RPC111-1D10.TPN RPCI-11 Homo sapiens genomic clone RPCI-11-ID10,
genomic survey sequence.
ACCESSION A0003188
VERSION A0003188.1 GI:3030392
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 110)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venter,J.C.
TITLE Use of BAC End Sequences for Sequence-Ready Map Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

FEATURES Location/Qualifiers
source 1..110
/organism="Homo sapiens"
/db_xref="GDB:7500081"
/db_xref="taxon:9606"
/clone="RPCI-11-ID10"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC Library"
BASE COUNT 22 a 27 c 26 g 35 t

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 110)
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 24, 1995 this sequence version replaced gi:634306.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1350
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuauqui,
M.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html

Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 90.

FEATURES

source
Location/Qualifiers

1..110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1007406"
/clone_lib="NCI-CGAP_Prl"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"

/note="Vector: PAMP10; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected, histologically normal
prostate epithelial cells. Double-stranded cDNA was
ligated to EcoRI adaptors, 5 cycles of PCR applied to the
cDNA with an adaptor-specific primer, and the resulting
PCR product subcloned into PAMP10 by the UDG-cloning
method (Life Technologies). Average insert size is 600
bp. NOTE: Not directionally cloned. This library was
constructed by David Krizman."

BASE COUNT 17 a 26 c 28 g 38 t 1 others

Query Match 0.3%; Score 86.6; DB 30; Length 110;
Best Local Similarity 86.4%; Pred. No. 0.38;
Matches 95; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 17609 TTTTCTTTTGAATAGATCTCGCTCTGTACCCAGGCTGAGTGCAGTGGCGCAATCT 17668

Db 1 TTTTCTTTTGTAGATGAGTCTGTGATCTGTGCCAGGCTGGAGTGCAGTGCGAGANTCT 60

QY 17669 CAGCTCAGTCGACGCGCCCTCGTGGGTTCAAGTATTCCTCGCTCA 17718

Db 61 TGGCTCATGCAACCTCTGCCCTCGTGGGTTCAAGAGATTCTCTCGCTCA 110

RESULT 13

H67040/c
LOCUS H67040 107 bp mRNA EST 27-OCT-1995
DEFINITION yu68c01.r1 Weizmann Olfactory Epithelium Homo sapiens cDNA clone
IMAGE:238944 5' similar to contains Alu repetitive element.; mRNA
sequence.

ACCESSION H67040

VERSION H67040.1 GI:1025780

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 107)

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chissoe, S., Dietrich, N., Dubuque, T., Pavello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

97044478

On Nov 29, 1993 this sequence version replaced gi:429999.

JOURNAL

MEDLINE

COMMENT

Contact: Willson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence stops: 101

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: M13RPI

High quality sequence stop: 101.

FEATURES

source

Location/Qualifiers

1..107

/organism="Homo sapiens"

/db_xref="GDB:3864328"

/db_xref="taxon:9606"

/clone="IMAGE:238944"

/clone_lib="Weizmann Olfactory Epithelium"

/sex="Female"

/tissue_type="olfactory epithelium"

/dev_stage="35 year old"

/lab_host="SOLR cells (kanamycin resistant)"

/note="Organ: nose; Vector: pBluescript SK-; Site_1:

EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:

Oligo dT. Olfactory epithelium, normal. Average insert

size: 0.8 kb; Uni-ZAP XR Vector. Library constructed by N.

Walker, D. lancet, Weizmann Institute of Science. -5'

adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor

sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'

BASE COUNT 24 a 37 c 20 g 24 t 2 others

Query Match 0.3%; Score 85.8; DB 24; Length 107;
Best Local Similarity 86.9%; Pred. No. 0.48;
Matches 93; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 10997 TAGCTGGCGTGGTGACATGCTCTAGTCCAGCTACTGGGAGGCTGAGCAGAGA 11056

Db 107 TAGCTGGGTTGTGTAGACATGCTCTGTATTCNAGCTACTCAGNAGGCTGAGTAGAGA 48

QY 11057 ATTGCTTGAACCTCGGAGCGGAGTTGCAGTCAGCGCGAGATTGCGC 11103

Db 47 ATCGCTTGAACCCAGGCGAGGTTGCAGTCAGCTGAGATTGTGC 1

RESULT 14

B65160/c

LOCUS B65160 108 bp DNA GSS 21-JUN-1998

DEFINITION CIT-HSP-2017G2.TRB CIT-HSP Homo sapiens genomic clone 2017G2,
genomic survey sequence.

ACCESSION B65160

VERSION B65160

KEYWORDS GSS.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 108)

AUTHORS Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K.,

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2000, 04:56:24 ; Search time 372.51 Seconds
(without alignments)
10119.719 Million cell updates/sec

Title: US-08-852-495c-1_copy_56000_85000
Perfect score: 29001
Sequence: 1 ATGACCAAGGCTGACTGAT.....CAGGAGACTAGAGTTTATT 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A-COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B-COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C-COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D-COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6-COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS-COMB.seq:*
7: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83.2	0.3	105	4	US-08-481-658B-65
2	83.2	0.3	105	4	US-08-477-504A-65
3	83.2	0.3	105	4	US-08-486-756A-65
4	83.2	0.3	105	4	US-08-485-862B-65
5	83.2	0.3	105	5	US-08-787-739-65
6	81.6	0.3	105	4	US-08-481-658B-65
7	81.6	0.3	105	4	US-08-477-504A-65
8	81.6	0.3	105	4	US-08-486-756A-65
9	81.6	0.3	105	4	US-08-485-862B-65
10	81.6	0.3	105	5	US-08-787-739-65
11	67.2	0.2	84	3	US-08-454-557C-91
12	67.2	0.2	84	4	US-08-340-426D-91
13	67.2	0.2	84	4	US-08-450-673C-91
14	67.2	0.2	84	6	PCT-US95-1711A-91
15	62.2	0.2	84	3	US-08-454-557C-91
16	62.2	0.2	84	4	US-08-340-426D-91
17	62.2	0.2	84	4	US-08-450-673C-91
18	62.2	0.2	84	6	PCT-US95-1711A-91
19	58.8	0.2	78	3	US-08-454-557C-70
20	58.8	0.2	78	4	US-08-340-426D-70
21	58.8	0.2	78	4	US-08-450-673C-70
22	58.8	0.2	78	6	PCT-US95-1711A-70
23	54.6	0.2	85	3	US-08-454-557C-92
24	54.4	0.2	85	3	US-08-454-557C-92
25	54.6	0.2	85	4	US-08-340-426D-92
26	54.4	0.2	85	4	US-08-340-426D-92
27	54.6	0.2	85	4	US-08-450-673C-92

Sequence 92, Appl
Sequence 92, Appl
Sequence 92, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 60, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 69, Appl
Sequence 66, Appl
Sequence 66, Appl
Sequence 66, Appl
Sequence 66, Appl
Sequence 66, Appl
Sequence 70, Appl
Sequence 70, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-65
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; Zip: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 13-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.3%; Score 83.2; DB 4; Length 105;
Best Local Similarity 87.5%; Pred. No. 1.3e-09;

Matches 91; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 27998 TTTTGTATTTTATTAGACAGGGTTTCACTATGTTGGCCAGGCTGATCTCAAACTCC 28057

Db 2 TTTTACATCTTTAGTAGACAGAGGTTTACACATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy 28058 TGACCTCATGATCCCGCTGCGCTTGGCCCTCTCAAAAGTGTGGGAT 28101

Db 62 TGACCTGTGTATCCACAGCCTCGGCCCTCCCAAAAGTGTGGGAT 105

RESULT 2

US-08-477-504A-65

; Sequence 65, Application US/08477504A

; Patent No. 5972353

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA: US/08/477,504A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3D

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-477-504A-65

Query Match 0.3%; Score 83.2; DB 4; Length 105;
Best Local Similarity 87.5%; Pred. No. 1.3e-09;
Matches 91; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 27998 TTTTGTATTTTATTAGACAGGGTTTCACTATGTTGGCCAGGCTGATCTCAAACTCC 28057

Db 2 TTTTACATCTTTAGTAGACAGAGGTTTACACATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy 28058 TGACCTCATGATCCCGCTGCGCTTGGCCCTCTCAAAAGTGTGGGAT 28101

Db 62 TGACCTGTGTATCCACAGCCTCGGCCCTCCCAAAAGTGTGGGAT 105

RESULT 3

US-08-486-756A-65

; Sequence 65, Application US/08486756A

; Patent No. 5981711

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA: US/08/486,756A

; APPLICATION NUMBER: US/08/486,756A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-486-756A-65

Query Match 0.3%; Score 83.2; DB 4; Length 105;
Best Local Similarity 87.5%; Pred. No. 1.3e-09;
Matches 91; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 27998 TTTTGTATTTTATTAGACAGGGTTTCACTATGTTGGCCAGGCTGATCTCAAACTCC 28057

Db 2 TTTTACATCTTTAGTAGACAGAGGTTTACACATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy 28058 TGACCTCATGATCCCGCTGCGCTTGGCCCTCTCAAAAGTGTGGGAT 28101

Db 62 TGACCTGTGTATCCACAGCCTCGGCCCTCCCAAAAGTGTGGGAT 105

RESULT 4

US-08-485-862B-65

; Sequence 65, Application US/08485862B

; Patent No. 5989938

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

;; CITY: Tiburon
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94920
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/485,862B
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/477,504
;; FILING DATE: 07-JUN-1995
;; APPLICATION NUMBER: US 08/260,190
;; FILING DATE: 15-JUN-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lauder, Leona L.
;; REFERENCE/DOCKET NUMBER: D-0021.3D
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-435-2034
;; TELEFAX: 415-435-0727
;; INFORMATION FOR SEQ ID NO: 65:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 105 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-08-485-862B-65

Query Match 0.3%; Score 83.2; DB 4; Length 105;
Best Local Similarity 87.5%; Pred. No. 1.3e-09;
Matches 91; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 27998 TTTTGTATTTTATTAGACAGGTTTCACACTGTTGGCCAGGCTGATCTCAAACTCC 28057
Db 2 TTTTACATCTTTAGTAGACAGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

QY 28058 TCACCTCATGATCCGCCCTGCCCTCGCCCTCAAGTGTGGGAT 28101
Db 62 TGACCTTGTGATCCACGCGCTCGGCCCTCCCAAGTGTGGGAT 105

RESULT 5
US-08-787-739-65
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/787,739
;; FILING DATE: 24-JAN-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/485,049
;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/486,756
;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/477,504
;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/481,658
;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/485,862
;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/485,863
;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/487,077
;; FILING DATE: 07-JUN-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lauder, Leona L.
;; REFERENCE/DOCKET NUMBER: 30,863
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-981-2034
;; TELEFAX: 415-981-0332
;; INFORMATION FOR SEQ ID NO: 65:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 105 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-08-787-739-65

Query Match 0.3%; Score 83.2; DB 5; Length 105;
Best Local Similarity 87.5%; Pred. No. 1.3e-09;
Matches 91; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 27998 TTTTGTATTTTATTAGACAGGTTTCACACTGTTGGCCAGGCTGATCTCAAACTCC 28057
Db 2 TTTTACATCTTTAGTAGACAGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

QY 28058 TCACCTCATGATCCGCCCTGCCCTCGCCCTCAAGTGTGGGAT 28101
Db 62 TGACCTTGTGATCCACGCGCTCGGCCCTCCCAAGTGTGGGAT 105

RESULT 6
US-08-481-658B-65/c
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA: US/08/481.658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.3%; Score 81.6; DB 4; Length 105;
Best Local Similarity 86.5%; Pred. No. 3e-09;
Matches 90; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 25018 ATCCGAGCTCTTTGGAGGCGCTAGCGGTGATCAGGAGTTCAGACGAGC 25077
Db 105 ATCCGAGCACTTTGGAGGCGCGAGGCTGGTGCATCACAAGGTTCAGGAGTTTGAGAGCAGC 46
Qy 25078 CTCGCCAAGATGGTGAATCCGCTCTACTATAAAGTATAAAA 25121
Db 45 CTGGCCATATGGTGAAACCCCTGCTCTACTAAGATGTAAAAA 2

RESULT 7
US-08-477-504A-65/c
Sequence 65, Application US/08477504A
Patent No. 5972353
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477.504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-477-504A-65
Query Match 0.3%; Score 81.6; DB 4; Length 105;
Best Local Similarity 86.5%; Pred. No. 3e-09;
Matches 90; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 25018 ATCCGAGCTCTTTGGAGGCGCTAGCGGTGATCAGGAGTTCAGACGAGC 25077
Db 105 ATCCGAGCACTTTGGAGGCGCGAGGCTGGTGCATCACAAGGTTCAGGAGTTTGAGAGCAGC 46
Qy 25078 CTCGCCAAGATGGTGAATCCGCTCTACTATAAAGTATAAAA 25121
Db 45 CTGGCCATATGGTGAAACCCCTGCTCTACTAAGATGTAAAAA 2

RESULT 8
US-08-486-756A-65/c
Sequence 65, Application US/08486756A
Patent No. 5981711
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486.756A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear


```
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-486-756A-65

Query Match          0.3%; Score 81.6; DB 4; Length 105;
Best Local Similarity 86.5%; Pred. No. 3e-09;
Matches 90; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 25018 ATCCGAGCTCTTTGGAGGCGCTAGGCGGTGGATCAGAGGTCAGAGTTCAAGACCAGC 25077
||||| ||||||| ||||||| ||| ||||||| ||||||| ||||||| ||| ||||
Db 105 ATCCGAGCACTTTGGAGGCGCGGCTGGTGGATCACAAGTCAGAGTTTGAGACGAGC 46
||||| ||||||| ||||||| ||| ||||||| ||||||| ||||||| ||| ||||

QY 25078 CTCGCCAAGATGTTGAATCCCGTCTCTACTATAAAGTATATAAAA 25121
|| ||||| ||||||| || ||||||| || ||||||| || |||||||
Db 45 CTGGCCAATATGTTGAACCCCTGCTCTACTATAAAGATGTAATAA 2

RESULT 10
US-08-787-739-65/c
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-787-739-65

Query Match          0.3%; Score 81.6; DB 4; Length 105;
Best Local Similarity 86.5%; Pred. No. 3e-09;
Matches 90; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 25018 ATCCGAGCTCTTTGGAGGCGCTAGGCGGTGGATCAGAGGTCAGAGTTCAAGACCAGC 25077
||||| ||||||| ||||||| ||| ||||||| ||||||| ||||||| ||| ||||
Db 105 ATCCGAGCACTTTGGAGGCGCGGCTGGTGGATCACAAGTCAGAGTTTGAGACGAGC 46
||||| ||||||| ||||||| ||| ||||||| ||||||| ||||||| ||| ||||

QY 25078 CTCGCCAAGATGTTGAATCCCGTCTCTACTATAAAGTATATAAAA 25121
|| ||||| ||||||| || ||||||| || ||||||| || |||||||
Db 45 CTGGCCAATATGTTGAACCCCTGCTCTACTATAAAGATGTAATAA 2

RESULT 9
US-08-485-862B-65/c
; Sequence 65, Application US/08485862B
; Patent No. 5989838
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,862B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-485-862B-65

Query Match          0.3%; Score 81.6; DB 4; Length 105;
Best Local Similarity 86.5%; Pred. No. 3e-09;
Matches 90; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
```

Query Match 0.3%; Score 81.6; DB 5; Length 105;
Best Local Similarity 86.5%; Pred. No. 3e-09;
Matches 90; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 25018 ATCCAGCCTTTGGAGGCTAGCGGTGATCAGGAGTTCAGACACAGC 25077
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 105 ATCCAGCAGCTTTGGAGGCGCGAGCTGGTGCATCACAAGTTCAGGAGTTCAGAGACAGC 46
QY 25078 CTCGCCAAGATGGTGAATCCGCTCTCTACTAAAGTATATAAAA 25121
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 45 CTGGCCAATATGGTGAACCCCTGTCTCTACTAAAGATGTAATAA 2

RESULT 11
US-08-454-557C-91
; Sequence 91, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454.557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-454-557C-91

Query Match 0.2%; Score 67.2; DB 3; Length 84;
Best Local Similarity 90.0%; Pred. No. 3.8e-06;
Matches 72; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 11636 TGTAAACAGGATGCTCTGATCTCTGACCTCGTGATCGGCCCGCCCTCAGCCTCCCAAA 11695
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4 TGTTCATCAGGCTGGTGTGCAACTCTGACCTCGTGATCGGCCCGCCCTCAGCCTCCCAAA 63
QY 11696 GTGCTGGGATTACAGGAGTG 11715
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 64 GTGCTGGGATTACAAAGCGTG 83

RESULT 12
US-08-340-426D-91
; Sequence 91, Application US/08340426D
; Patent No. 5948634

GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340.426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-340-426D-91

Query Match 0.2%; Score 67.2; DB 4; Length 84;
Best Local Similarity 90.0%; Pred. No. 3.8e-06;
Matches 72; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 11636 TGTAAACAGGATGCTCTGATCTCTGACCTCGTGATCGGCCCGCCCTCAGCCTCCCAAA 11695
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4 TGTTCATCAGGCTGGTGTGCAACTCTGACCTCGTGATCGGCCCGCCCTCAGCCTCCCAAA 63
QY 11696 GTGCTGGGATTACAGGAGTG 11715
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 64 GTGCTGGGATTACAAAGCGTG 83

RESULT 13
US-08-450-673C-91
; Sequence 91, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

Search completed: June 15, 2000, 13:18:30
Job time: 87840 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 15, 2000, 13:12:08 ; Search time 17972.2 Seconds
(without alignments)
-1569.757 Million cell updates/sec

Title: US-08-852-495c-l_copy_84000_113000
Perfect score: 29001
Sequence: 1 TTGTATTTTATTAGAGACA.....GCCTCAGCCTTCAGAGTGTA 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_om:*
- 4: gb_ov:*
- 5: gb_pat:*
- 6: gb_ph:*
- 7: gb_pl1:*
- 8: gb_pl2:*
- 9: gb_pl3:*
- 10: gb_pr2:*
- 11: gb_pr3:*
- 12: gb_ro:*
- 13: gb_sts:*
- 14: gb_sy:*
- 15: gb_un:*
- 16: gb_vl:*
- 17: em_fun:*
- 18: em_hum1:*
- 19: em_hum2:*
- 20: em_in:*
- 21: em_om:*
- 22: em_or:*
- 23: em_ov:*
- 24: em_pat:*
- 25: em_ph:*
- 26: em_pl:*
- 27: em_ro:*
- 28: em_sts:*
- 29: em_sy:*
- 30: em_un:*
- 31: em_vl:*
- 32: gb_htg1:*
- 33: gb_htg2:*
- 34: gb_in1:*
- 35: gb_in2:*
- 36: em_ba1:*
- 37: em_ba2:*
- 38: em_hum3:*
- 39: em_hum4:*
- 40: gb_pr4:*
- 41: gb_htg3:*
- 42: gb_htg4:*
- 43: gb_htg5:*
- 44: gb_htg6:*

- 45: gb_htg7:*
- 46: em_htg1:*
- 47: em_htg2:*
- 48: em_htg3:*
- 49: em_hums:*
- 50: gb_pl3:*
- 51: gb_pr5:*
- 52: gb_htg8:*
- 53: gb_htg9:*
- 54: gb_htg10:*
- 55: gb_htg11:*
- 56: gb_htg12:*
- 57: gb_htg13:*
- 58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	95	0.3	107	9	HUMALCE162	M87924 Human carci
2	89	0.3	108	11	HSU67803	U67803 Human small
3	87.4	0.3	108	10	HSLDLRN2	X05250 Human LDL-r
4	86.8	0.3	108	10	HSLDLRN2	X05250 Human LDL-r
5	85.4	0.3	107	9	HUMALCE162	M87924 Human carci
6	83	0.3	103	9	HUMALCE221	M87896 Human carci
7	81.4	0.3	103	9	HUMALCE221	M87896 Human carci
8	81	0.3	108	11	HSU67803	U67803 Human small
9	80	0.3	108	10	HSLDLRD1	X05249 Human LDL-r
10	80	0.3	108	10	HSLDLRD2	X05251 Human LDL-r
11	79.4	0.3	108	10	HSLDLRD1	X05249 Human LDL-r
12	79.4	0.3	108	10	HSLDLRD2	X05251 Human LDL-r
13	78.2	0.3	108	10	HSLDLI12	X05248 Human LDL-r
14	77.8	0.3	108	11	HSU67804	U67804 Human small
15	76.4	0.3	108	11	HSU67808	U67808 Human small
16	76.4	0.3	110	9	HUMALCE43	M87900 Human carci
17	76	0.3	104	9	HUMALCE272	M87899 Human carci
18	75.4	0.3	97	9	HUMDLIRA2	M1480 Human low d
19	74	0.3	110	9	HUMALCE43	M87900 Human carci
20	73	0.3	108	11	HSU67808	U67808 Human small
21	72.6	0.3	106	13	G32743	G32743 A009P31 Hum
22	72.8	0.3	110	11	HSU67807	U67807 Human small
23	72.2	0.2	108	9	HUMD1D03M5	D16965 Human HepG2
24	72.4	0.2	108	10	HSLDLI12	X05248 Human LDL-r
25	72	0.2	90	9	HUMDLRFL	K03555 Human low d
26	71.2	0.2	103	13	HS8IC8R	X57789 Human sequ
27	71.4	0.2	108	11	HSU67804	U67804 Human small
28	70.8	0.2	91	13	HUMUT8164A	L30244 Human STS U
29	70.4	0.2	104	9	HUMALCE272	M87899 Human carci
30	70	0.2	107	11	HSU67806	U67806 Human small
31	69.4	0.2	110	11	HSU67807	U67807 Human small
32	68.8	0.2	80	9	HUMBRKFAE	M36135 Human alpha
33	68.2	0.2	107	11	HSU67806	U67806 Human small
34	67.2	0.2	80	9	HUMBRKFAE	M36135 Human alpha
35	67.4	0.2	94	9	HUMHGAL	M13479 Human alpha
36	67.4	0.2	95	13	HUMUT8002B	L30176 Human STS U
37	66.6	0.2	91	13	HUMUT8164A	L30244 Human STS U
38	66.8	0.2	94	9	HUMHGAL	M13479 Human alpha
39	66.8	0.2	95	13	HUMUT8002B	L30176 Human STS U
40	67	0.2	103	13	HS8IC8R	X57789 Human sequ
41	66.2	0.2	108	13	G43555	G43555 WIAP-2393-S
42	65.6	0.2	97	9	HUMDLRDJ	M14179 Human famli
43	65.2	0.2	79	10	S73203	S73203 ALL-1 (tand
44	65.2	0.2	108	9	HUMD1D03M5	D16965 Human HepG2
45	64.8	0.2	76	11	AF032287	AF032287 Eulemur m

ALIGNMENTS

```

RESULT 1
HUMALCE162      107 bp ss-RNA      PRI      15-APR-1994
LOCUS           Human carcinoma cell-derived Alu RNA transcript, clone CE162.
DEFINITION
ACCESSION      M87924
VERSION        M87924.1 GI:174871
KEYWORDS       Alu repeat.
SOURCE         Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS       Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE         Alu RNA transcripts in human embryonal carcinoma cells. Model of
              post-transcriptional selection of master sequences
JOURNAL       J. Mol. Biol. (1992) In press
FEATURES      Location/Qualifiers
               1..107
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /cell_line="NTera2D1"
               /dev_stage="embryo"
               /sex="male"
               /tissue_type="carcinoma"
BASE COUNT    28 a 30 c 35 g 14 t
ORIGIN

Query Match    0.3%; Score 95; DB 9; Length 107;
Best Local Similarity 95.1%; Pred. No. 1e-08;
Matches 98; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7316 GGCAGAGAAATGGCTGACCGGGAGCGAGCTTGACGTGAGCGGAGATCGGCCCATG 7375
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 5 GGCAGAGAAATGGCTGACCGGGAGCGAGCTTGACGTGAGCGGAGATCGGCCCATG 64

QY 7376 GCACCTCCAGCCTGGGTGACAGCGGAGACTCGTCTCAAAAAA 7418
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65 GCACCTCCAGCCTGGGTGACAGCGGAGACTCGTCTCAAAAAA 107

RESULT 2
HSU67803
LOCUS           HSU67803 108 bp RNA      PRI      01-AUG-1997
DEFINITION      Human small cytoplasmic Alu transcript.
ACCESSION      U67803
VERSION        U67803.1 GI:2289917
KEYWORDS       Alu.
SOURCE         human.
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS       Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE         Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
              cDNAs derived from primary and small cytoplasmic Alu (scAlu)
              transcripts
JOURNAL       J. Mol. Biol. 271 (2), 222-234 (1997)
MEDLINE       97415756
REFERENCE      2 (bases 1 to 108)
AUTHORS       Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE         Direct Submission
JOURNAL       Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
              Children's Hospital of Philadelphia, 1004F Abramson Research
              Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES      Location/Qualifiers
               1..108
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="TscAlu2"
repeat_region  1..108
               /note="scAlu"
               /rpt_family="Alu"

```

```

BASE COUNT    23 a 39 c 30 g 16 t
ORIGIN

Query Match    0.3%; Score 89; DB 11; Length 108;
Best Local Similarity 94.8%; Pred. No. 1.5e-07;
Matches 92; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2770 GCCTGTAATCTAGACATTTGGGAGCGGAGCGGCGGATCACGAGGTCAGGATCGA 2829
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 GCCTGTAATCTAGACATTTGGGAGCGGAGCGGCGGATCACGAGGTCAGGATCGA 60

QY 2830 GACCATCTTGGTACACGCGTGAACCCCGTTCTTAC 2866
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 GACCATCTTGGTACACGCGTGAACCCCGTTCTTAC 97

RESULT 3
HSLDLRN2/c
LOCUS           HSLDLRN2 108 bp DNA      PRI      20-MAY-1992
DEFINITION      Human LDL-receptor gene Intron 14 fragment (normal gene).
ACCESSION      X05250
VERSION        X05250.1 GI:34337
KEYWORDS       Alu repetitive sequence; low density lipoprotein receptor.
SOURCE         human.
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
AUTHORS       Primates; Catarrhini; Hominiidae; Homo.
              1 (bases 1 to 108)
              Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
              Williamson,R. and Humphries,S.
              Unequal crossing-over between two alu-repetitive DNA sequences in
              the low-density-lipoprotein-receptor gene. A possible mechanism for
              the defect in a patient with familial hypercholesterolaemia
              Eur. J. Biochem. 164 (1), 77-81 (1987)
JOURNAL       87161901
MEDLINE
COMMENT        See X05252 for deletion junction
              Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES      Location/Qualifiers
               1..108
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /note="intron XIV fragment"
BASE COUNT    28 a 23 c 39 g 18 t
ORIGIN

Query Match    0.3%; Score 87.4; DB 10; Length 108;
Best Local Similarity 89.5%; Pred. No. 3.1e-07;
Matches 94; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 9676 CTCGGCTCAGCGCAACCTCCGCTCCAGGGTTCAAGCAATTCCTCGCTCAGCTCCCG 9735
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 108 CTCGGCTCAGTCAACCTCTGCTCTCGCTTCAAGCAATTCCTCGCTCAGCTCCCG 49

QY 9736 ACTAATCGGACTACTGCGAAGCGCCACCGCTGGCTAATTTT 9780
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 48 AGTAGCTGGATTACAGGCACTGCCACACGCTGGCTAATTTT 4

RESULT 4
HSLDLRN2
LOCUS           HSLDLRN2 108 bp DNA      PRI      20-MAY-1992
DEFINITION      Human LDL-receptor gene Intron 14 fragment (normal gene).
ACCESSION      X05250
VERSION        X05250.1 GI:34337
KEYWORDS       Alu repetitive sequence; low density lipoprotein receptor.
SOURCE         human.
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Chordata; Vertebrata; Mammalia; Eutheria;
AUTHORS       Primates; Catarrhini; Hominiidae; Homo.

```


Query Match 0.3%; Score 80; DB 10; Length 108;
Best Local Similarity 85.6%; Pred. No. 8.8e-06;
Matches 89; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 9677 TCGGCTACCCGACACCTCCGCTCCAGGTTCAAGCAATTCCTCGCTCAGCTCCGCCA 9736

Db 107 TCGGCTACCCACACCTCGCTCCTCGGTTCAACACCAATTTTCTGCTCAGCTCCGCCA 48

Qy 9737 GTAATTGGGACTTGGCAAGCGCACCCAGCTGGCTAAATTTT 9780

Db 47 GTAGCTGGGATTACAGGCACCTGCCACCACCGCTGGCTAAATTTT 4

RESULT 11

HSLLDLR1/c 108 bp DNA PRI 20-MAY-1992
LOCUS Human LDL-receptor mutated gene with intron 12 deletion junction.
DEFINITION X05249

ACCESSION X05249.1 GI:34335

VERSION X05249.1

KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 108)

AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,

TITLE Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia

Eur. J. Biochem. 164 (1), 77-81 (1987)

JOURNAL 87161901

MEDLINE

COMMENT

*source: hypercholesterol aemia
See X05248 for corresponding normal gene sequence
In the defective LDL-receptor gene the deletion occurred between two alu-repetitive sequences, that are in the same direction, the deletion eliminates exons 13 and 14 and changes the reading frame of the resulting spliced mRNA.

Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.

FEATURES

source

1. .108

/organism="Homo sapiens"

/db_xref="taxon:9606"

/cell_type="blood leukocytes from a patient with familial"

misc_feature

1. .108

/note="deletion junction region intron 12/ intron 15"

BASE COUNT 20 a 40 c 20 g 28 t

ORIGIN

Query Match 0.3%; Score 79.4; DB 10; Length 108;

Best Local Similarity 84.8%; Pred. No. 1.1e-05;

Matches 89; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 2877 AAAAATTAGCGGGCGTGTGGCGGGCCCTGTAGTCCAGCTACTTGGAGGCTGAGGC 2936

Db 106 AAAATTAGCAGGCGTGTGGCAGGTCCTGTATCCAGCTACTCGGAGGCTGAGGC 47

Qy 2937 AGGAGAATGGCATGAACCTGGAGCGGAGCTTGCGAGTGAGCCGA 2981

Db 46 AGGAAATGGTTTGAACCCAGGAGGAGGTTGTGGTGAGCGCA 2

RESULT 12

HSLLDLR2

LOCUS Human LDL-receptor mutated gene with intron 14 deletion junction.

DEFINITION X05251

ACCESSION X05251.1 GI:34336

VERSION X05251.1

KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 108)

AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,

Williamson,R. and Humphries,S.

TITLE Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia

Eur. J. Biochem. 164 (1), 77-81 (1987)

JOURNAL 87161901

MEDLINE

COMMENT

*source: hypercholesterol aemia

See X05250 for corresponding normal gene sequence
In the defective LDL-receptor gene the deletion occurred between two alu-repetitive sequences, that are in the same direction, the deletion eliminates exons 13 and 14 and changes the reading frame of the resulting spliced mRNA.

Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.

FEATURES

source

1. .108

/organism="Homo sapiens"

/db_xref="taxon:9606"

/cell_type="blood leukocytes from a patient with familial"

intron

BASE COUNT 28 a 20 c 40 g 20 t

ORIGIN

Query Match 0.3%; Score 79.4; DB 10; Length 108;

Best Local Similarity 84.8%; Pred. No. 1.1e-05;

Matches 89; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 2877 AAAAATTAGCGGGCGTGTGGCGGGCCCTGTAGTCCAGCTACTTGGAGGCTGAGGC 2936

Db 3 AAAAATTAGCAGGCGTGTGGCAGGTCCTGTATCCAGCTACTCGGAGGCTGAGGC 62

Qy 2937 AGGAGAATGGCATGAACCTGGAGCGGAGCTTGCGAGTGAGCCGA 2981

Db 63 AGGAAATGGTTTGAACCCAGGAGGAGGTTGTGGTGAGCGCA 107

RESULT 13

HSLLDL12

LOCUS Human LDL-receptor gene intron 12 fragment (normal gene) LDL - low

DEFINITION density lipoprotein.

ACCESSION X05248

VERSION X05248.1 GI:34334

KEYWORDS Alu repetitive sequence; low density lipoprotein receptor;

repetitive sequence.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 108)

AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,

Williamson,R. and Humphries,S.

TITLE Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia

Eur. J. Biochem. 164 (1), 77-81 (1987)

JOURNAL 87161901

MEDLINE

COMMENT

*source: hypercholesterol aemia

See X05249 for deletion junction

Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.

FEATURES

source

1. .108

/organism="Homo sapiens"

/db_xref="taxon:9606"

complement(<1. .65)

/note="Alu repeat"

intron

BASE COUNT 21 a 38 c 20 g 29 t

ORIGIN

Query Match	0.3%;	Score 78.2;	DB 10;	Length 108;
Best Local Similarity	83.2%;	Pred. No. 2e-05;		
Matches	89;	Conservative	0;	Mismatches 18; Indels 0; Gaps 0;
<hr/>				
QY	9677	TCGGCTCACCGCAACCTCCGGCTCCAGGGTTCAAGCAATTTCTCTGGCTCAGCCTCCCA	9736	
Db	2	TCGGCTCACCAACACCTCTGCTCTGGTTCAAAACCATTTCTCTGGCTCAGCCTCCCTTA	61	
<hr/>				
QY	9737	GTAATTCGGACTACTGCAAGCGCCACCCAGCCCTGGCTAATTTGTA	9783	
Db	62	GTAGCTGGGATTACAAGCATGTGCCACACGCCCGGCTGATTTGTA	108	
<hr/>				
RESULT	14			
LOCUS	HSU67804	108 bp	RNA	PRI 01-AUG-1997
DEFINITION	Human small cytoplasmic Alu transcript.			
ACCESSION	U67804			
VERSION	U67804.1	GI:2289918		
KEYWORDS	Alu.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 108)			
TITLE	Shaikh, T.H., Roy, A.M., Kim, J., Batzer, M.A. and Deininger, P.L. CDNA's derived from primary and small cytoplasmic Alu (scAlu) transcripts			
JOURNAL	J. Mol. Biol. 271 (2), 222-234 (1997)			
MEDLINE	97415756			
REFERENCE	2 (bases 1 to 108)			
AUTHORS	Shaikh, T.H., Kim, J., Batzer, M.A. and Deininger, P.L.			
TITLE	Direct Submission			
JOURNAL	Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The Children's Hospital of Philadelphia, 1004F Abramson Research Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA			
<hr/>				
FEATURES	Location/Qualifiers			
source	1..108			
repeat_region	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="TscAlu3" 1..108 /note="scAlu" /rpt_family="Alu" /rpt_type=dispersed			
BASE COUNT	26 a 38 c 26 g			18 t
ORIGIN				
<hr/>				
Query Match	0.3%;	Score 77.8;	DB 11;	Length 108;
Best Local Similarity	87.6%;	Pred. No. 2.4e-05;		
Matches	85;	Conservative	0;	Mismatches 12; Indels 0; Gaps 0;
<hr/>				
QY	2770	GCCTGTATCTAGCACTTTGGAGGCGGACAGCGGGATCACGAGTCAAGGATCGA	2829	
Db	1	GCCTGTATCCACACTTTGGAAGGCAAGCGGGAGGATCAAGTCAAGGATCGA	60	
<hr/>				
QY	2830	GACCATCTGGCTAACACGCTGAACCCCGTTTCTAC	2866	
Db	61	GACCATCTGGCTAACATGGTGAACCCCGCTTTCC	97	
<hr/>				
RESULT	15			
LOCUS	HSU67808/c	108 bp	RNA	PRI 01-AUG-1997
DEFINITION	Human small cytoplasmic Alu transcript.			
ACCESSION	U67808			
VERSION	U67808.1	GI:2289922		
KEYWORDS	Alu.			
SOURCE	human.			
ORGANISM	Homo sapiens			

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE
```

```
AUTHORS Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.  
TITLE cDNAs derived from primary and small cytoplasmic Alu (scAlu) transcripts  
  
JOURNAL J. Mol. Biol. 271 (2), 222-234 (1997)  
MEDLINE 97415756  
REFERENCE
```

```
AUTHORS Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.  
TITLE Direct Submission  
JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The Children's Hospital of Philadelphia, 1004F Abramson Research Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
```

```
FEATURES source 1..108 /organism='Homo sapiens' /db_xref='taxon:9606' /clone='TSCALU7'  
repeat_region 1..108 /note='scAlu' /rpt_family='Alu' /rpt_type=dispersed  
BASE COUNT 22 a 37 C 28 G 21 t  
ORIGIN
```

```
Query Match 0.3% Score 76.4 ; DB 11; Length 108;  
Best Local Similarity 88.3%; Pred. No. 4.4e-05;  
Matches 83; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```
QY 14 AGAGACAGGGTTTCACATGTTGTGCCAGCGTGACTCAAACTCCTGCACCTCATGATCCGC 73  
| | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 95 AAAGACGGAGTTTCACCATTGTGCCAGCGTTGGTCTCAAACCTCCTGTGATCCAC 36  
| | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
QY 74 CTGCCTTGGCCTCTCAAAGTAGCTGGGATTACAGG 107  
| | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 35 CCGACTTGGCCTCCCAGAAGTCTGGGATTACAGG 2  
| | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Search completed: June 16, 2000, 00:11:20  
Job time: 127964 sec
```

Search completed: June 16, 2000, 00:11:20
Job time: 127964 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: June 15, 2000, 16:12:36 ; Search time 593.87 seconds
(without alignments)
12217.860 Million cell updates/sec

Title: US-08-852-495C-1_COPY_84000_113000
Perfect score: 29001
Sequence: 1 TTGTAATTTTATTAGACAAAA.....GCCTCAGCCTTCAGAGTGA 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 433070

Minimum DB seq length: 10

Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	67.2	0.2	100	1 T24892	Human gene signatu
C 2	67	0.2	108	1 T26828	Human gene signatu
C 3	66.8	0.2	108	1 X12095	Human biallelic po
C 4	65.6	0.2	100	1 T24892	Human gene signatu
C 5	63.8	0.2	86	1 V41231	Mouse embryonic ce
C 6	62.8	0.2	103	1 T20927	Human gene signatu
C 7	61	0.2	108	1 T25009	Human gene signatu
C 8	60	0.2	93	1 T25688	Human gene signatu
C 9	58.6	0.2	108	1 T25009	Human gene signatu
C 10	58.6	0.2	110	1 T26288	Human gene signatu
C 11	58.4	0.2	110	1 T25260	Human gene signatu
C 12	57.2	0.2	108	1 T26828	Human gene signatu
C 13	57.2	0.2	108	1 X12095	Human biallelic po
C 14	56.2	0.2	103	1 T26213	Human gene signatu
C 15	54.8	0.2	97	1 T26728	Human gene signatu
C 16	55	0.2	100	1 X12087	Human biallelic po
C 17	55	0.2	100	1 X12085	Human biallelic po
C 18	55	0.2	100	1 X12086	Human biallelic po
C 19	54.6	0.2	69	1 Q29016	Probe to internal
C 20	54.6	0.2	91	1 T25054	Human gene signatu
C 21	54.4	0.2	109	1 T23895	Human gene signatu
C 22	53.8	0.2	99	1 T20931	Human gene signatu
C 23	53.4	0.2	97	1 T26728	Human gene signatu
C 24	53.6	0.2	100	1 Q75490	Human genome fragm
C 25	53	0.2	82	1 T25468	Human gene signatu
C 26	52	0.2	101	1 T24667	Human gene signatu
C 27	52.2	0.2	110	1 T26288	Human gene signatu
C 28	51.4	0.2	91	1 T25854	Human gene signatu
C 29	51.4	0.2	102	1 T20743	Human gene signatu
C 30	51	0.2	88	1 T21564	Human gene signatu
C 31	51	0.2	89	1 T23513	Human gene signatu
C 32	50.4	0.2	62	1 T25689	Human gene signatu
C 33	50.4	0.2	95	1 T23131	Human gene signatu
C 34	50.6	0.2	100	1 X12085	Human biallelic po

C 35	50.6	0.2	100	1 X12086	Human biallelic po
C 36	49.8	0.2	74	1 T25218	Human gene signatu
C 37	49.6	0.2	84	1 T25848	Human gene signatu
C 38	49.8	0.2	103	1 T26213	Human gene signatu
C 39	49.2	0.2	70	1 N60231	Normal chromosome
C 40	49.2	0.2	95	1 Q75099	Plasmid pOKSCI8a c
C 41	49.4	0.2	100	1 X12087	Human biallelic po
C 42	48.4	0.2	89	1 T23513	Human gene signatu
C 43	48.4	0.2	93	1 T22572	Human gene signatu
C 44	48	0.2	48	1 Q69409	Human H4/a gene fo
C 45	48	0.2	48	1 T63871	Human H4/a gene (f

ALIGNMENTS

RESULT 1
T24892/c
ID T24892 standard; cDNA to mRNA; 100 BP.
AC T24892;
DT 05-NOV-1996 (first entry)
DE Human gene signature HUMGS06998.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K. Okubo K;
PT WPI; 95-206931/27.
DR Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1720; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(I) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;

Query Match 0.2%; Score 67.2; DB 1; Length 100;
Best Local Similarity 78.8%; Pred. No. 0.0031;
Matches 78; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 8535 TTTTGTGTTTGTGAGCTGAGCCCTTGTCTGTACCCAGGCTGCGATGCGGATC 8594

Db 100 TTGTTGTTTTCACAGAGTGTACTCTGTACCCAGGCTGCGATGCGGATC 41

QY 8595 TCGGCTCACTGCAACCTCCGCTTCCAGGTTCAAGCGAT 8633

Db 40 TCAGCTNATGCAAAATCTGCCTCCCGAGGTTCAAGCGAT 2

RESULT 2

T26828/c
ID T26828 standard; cDNA to mRNA; 108 BP.

T26828;
14-NOV-1996 (first entry)
Human gene signature HUMG09078.
Gene signature; messenger RNA; mRNA; relative abundance; frequency;
human; cloning; mapping; non-biased library; diagnosis; detection;
cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PI Matsubara K, Okubo K;
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
for diagnosis of abnormal cell function, by preparing cDNA that
reflects relative abundance of corresp. mRNA in specific human
tissues
PS Claim 1; Page 2182; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
double-stranded DNA) which comprises one of the 7837 "GS" sequences
given in T19001-T26837 and which is able to hybridise to part of
human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
sequences were obtained from 3'-directed cDNA libraries prepared
from various human tissues; synthesis of cDNA was initiated from the
3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
untranslated sequence is unique to a particular mRNA species, almost
all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
is constructed so as to reflect accurately the relative abundance of
different mRNAs in the particular tissue from which it was derived.
The appearance frequency of a given GS in a cDNA library can be
determined (esp. using primers and probes derived from the GS
sequences) as a means of diagnosing abnormal cell function or for
recognising different cell types.
SQ Sequence 108 BP; 18 A; 33 C; 23 G; 28 T;
Query Match 0.2%; Score 67; DB 1; Length 108;
Best Local Similarity 90.9%; Pred. No. 0.0034;
Matches 70; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 2751 GCCGGCGCGTGGCTCAGCGCTGAATCCCTAGACTTTGGAGGCCGAGCGCGGAT 2810
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 77 GCCGGCGTGGTGGCTCATGCTGTAANCCAGCACTATGGGAGGCCGAGCGCGGAT 18
Qy 2811 CACGAGGTCAGGAGATC 2827
||||||| ||||||| |||||||
Db 17 GACGAGGTCAGGAGATC 1
RESULT 3
X12095
ID X12095 standard; DNA; 108 BP.
AC X12095;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment TIGR-A003M18a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
detection; phenotypic typing; characteristic; infection; hereditary;
autoimmune disease; cancer; inflammation; drug; therapy; medicament;
treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
determining polymorphic forms for use in e.g. forensics, paternity
testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.

CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
markers which have been isolated using the primers represented in
X09121-X10268. The base occupying the polymorphic site is indicated by
the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
methods for determining polymorphic forms in an individual for use in
e.g. forensics, paternity testing or for phenotypic typing for diseases
such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
hypercholesterolemia, polycystic kidney disease, hereditary
spherocytosis, von Willebrand's disease, tuberosus sclerosis, hereditary
haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
syndrome, osteogenesis imperfecta, acute intermittent porphyria,
autoimmune diseases, inflammation, cancer, diseases of the nervous
system, infection by pathogenic microorganisms, and characteristics such
as longevity, appearance (e.g. baldness, obesity), strength, speed,
endurance, fertility, and susceptibility or receptivity to particular
drugs or therapeutic treatments. The isolated polymorphic nucleic acid
segments can also be used to produce medicaments for the treatment or
prophylaxis of such diseases.
SQ Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;
Query Match 0.2%; Score 66.8; DB 1; Length 108;
Best Local Similarity 81.5%; Pred. No. 0.0037;
Matches 88; Conservative 1; Mismatches 18; Indels 1; Gaps 1;
Qy 9780 TGTATTTTATTAGAGATGAGGTTTCTCCATGTTGGTCAGACTGCTCGAACTCCCGAC 9839
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 TGTCTTTTGTAGAGATGAGGTTTCTCTGTGGCCAGGATGCTCGAACTCCTGAC 60
Qy 9840 CTCAGTGATGATCCACCGCTCGGCTCC-AAAAGTCTGGGATTACAG 9886
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 TTCAAGTCATCGTCTGCTGGCTCCCAAAAGTCTGGGATTATAG 108
RESULT 4
T24892
ID T24892 standard; cDNA to mRNA; 100 BP.
AC T24892;
DT 05-NOV-1996 (first entry)
DE Human gene signature HUMG06998.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
human; cloning; mapping; non-biased library; diagnosis; detection;
cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PI Matsubara K, Okubo K;
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
for diagnosis of abnormal cell function, by preparing cDNA that
reflects relative abundance of corresp. mRNA in specific human
tissues
PS Claim 1; Page 1720; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
double-stranded DNA) which comprises one of the 7837 "GS" sequences
given in T19001-T26837 and which is able to hybridise to part of
human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
sequences were obtained from 3'-directed cDNA libraries prepared
from various human tissues; synthesis of cDNA was initiated from the
3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
untranslated sequence is unique to a particular mRNA species, almost
all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
is constructed so as to reflect accurately the relative abundance of
different mRNAs in the particular tissue from which it was derived.
The appearance frequency of a given GS in a cDNA library can be
determined (esp. using primers and probes derived from the GS
sequences) as a means of diagnosing abnormal cell function or for
recognising different cell types.

SQL	Sequence	100 BP;	28 A;	22 C;	25 G;	22 T;				
	Query Match	0.2%;	Score 65.6;	DB 1;	Length 100;					
	Best Local Similarity	77.8%;	Pred. No. 0.0058;							
	Matches	77;	Conservative	0;	Mismatches	22;	Indels	0;	Gaps	0;
QY	2943	ATGCGATGAACCTGGGAGGCGGACCTTGCAGTGAGCGGAGATTCGGCAGCTGCACCTCCAA	3002							
DB	2	ATGCGTTGAACCTGGGAGGCGAATTTGCATNAGCTGAGATTCACCTTGCACCTCCG	61							
QY	3003	CTGTGGGAGACACGCGAGACTCCGCTCTCAAAAAA	3041							
DB	62	CTGTGGTGACAGAGTGACACTCTGTTGAAACAACAA	100							
RESULT	5									
ID	V41231									
AC	V41231	standard;	CDNA;	86 BP.						
DE	V41231;									
DT	01-OCT-1998	(first entry)								
DE	Mouse embryonic cell	EST 13-4 nucleotide sequence.								
DE	Embryonic stem cell;	ESC; non-primate;	mouse;	EST;	human;					
KW	developmental gene;	transgenic animal;	reporter gene;	ss.						
OS	Mus sp.									
PN	W09823633-A1.									
PD	04-JUN-1998.									
PF	25-NOV-1997;	U22335.								
PF	27-NOV-1996;	US-032510.								
PA	(CORR) CORNELL RES FOUND INC.									
PI	Holzschu DL, Mark WH;									
DR	WPI; 98-322656/28.									
PT	Screening for human developmental genes - by trapping in murine									
PT	embryonic stem cells and analysing differential expression in vitro,									
PT	selecting homologous non-human primate gene and using it to isolate									
PT	human gene									
PS	Claim 37; Page 18; 60pp; English.									
CC	Sequences shown in V41230 to V41247 represent nucleotide sequences of									
CC	mouse EST from tagged cDNA clones. These are used in the method of the									
CC	invention of screening for human developmental genes. The method									
CC	comprises inserting a promoterless reporter gene into a non-primate									
CC	mammalian embryonic stem cell (ESC) genome and identifying cellular									
CC	transcripts that encode the reporter gene product. Fragments of genes									
CC	encoding these transcripts are cloned and sequenced. A gene encoding a									
CC	transcript that includes unknown sequences is selected and expression									
CC	level of the gene encoding the transcript, or part of it, in different									
CC	cell types and/or different developmental stages is detected. A gene									
CC	showing differential expression is selected and expression levels of a									
CC	homologous non-human primate gene, in different cell types and/or at									
CC	different developmental stages, using the non-primate transcript as									
CC	probe is detected. A homologous gene having the same pattern of									
CC	differential expression is selected and the non-primate gene, or part of									
CC	it is used to identify the homologous human gene. The ESC transcripts									
CC	identified by this method are used to generate transgenic animals									
CC	selected from rats, hamsters, rabbits, dogs, pigs, horses, cows, monkey,									
CC	baboon or chimpanzee for study of gene function. The method provides									
CC	rapid and large scale screening for human developmental genes, and									
CC	eliminates the need to analyse reporter gene expression in embryos.									
SQL	Sequence	86 BP;	16 A;	28 C;	30 G;	12 T;				
	Query Match	0.2%;	Score 63.8;	DB 1;	Length 86;					
	Best Local Similarity	85.5%;	Pred. No. 0.012;							
	Matches	71;	Conservative	0;	Mismatches	12;	Indels	0;	Gaps	0;
QY	16950	CTCTTGGGGCGGTGTGACCATCGCGAGGGTGGCGCTTTGGCTTAATATTCAGGCGGTGCTG	17009							
DB	2	CTGTGGGCGCGTGTGACCATCGCGAGGGCGCGTCTGTGCCAATATCCAGGCGGTGCTG	61							
QY	17010	CTGCCTAAGAAAACCTG								

PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 1748; 2245pp; Japanese.

CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

SQ Sequence 108 BP; 34 A; 31 C; 26 G; 15 T;

Query Match 0.2%; Score 61; DB 1; Length 108;
Best Local Similarity 74.5%; Pred. No. 0.036;
Matches 76; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 8532 TTTTGTGTTTGTGACACTGAGCCTTCTGCTGTCACCCAGCGTGGAGTGCATGGCGCG 8591
DB 103 TGGTGTGTTGTTTCAACAGGGCTTGTGCTGTGTCACCTAGGCTGGAATNACAGTGGCGTG 44

QY 8592 ATCTCGGCTCACTCAACCTCCGCTTCCAGGTTCAAGCGAT 8633
DB 43 ACCATGGCTCACTGACGCTTGGCTCATGGCTCAGCGCAT 2

RESULT 8
T25688/c ID T25688 standard; cDNA to mRNA; 93 BP.
AC T25688;
DT 09-OCT-1996 (first entry)
DE Human gene signature HUMGS07887.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-AL.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 1907; 2245pp; Japanese.

CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for

CC recognising different cell types.
SQ Sequence 93 BP; 25 A; 27 C; 24 G; 17 T;

Query Match 0.2%; Score 60; DB 1; Length 93;
Best Local Similarity 78.3%; Pred. No. 0.052;
Matches 72; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 11115 TTTGAGATGGAATTTCACTCTTGTGTCAGGCGGGTGCAGTGGCACAGTCTCAGCTC 11174
DB 93 TTTGAGATGGGCTTCACTCTTGTGTCACCCAGACTGGAGTGCAGTGGTGCATCAGCTC 34

QY 11175 ACTGCAACCTCCGCTCTCTGGGTTCAAGGAT 11206
DB 33 ACTGTGCTTCGGCTTCTCTGGGCTCAAGAGAT 2

RESULT 9
T25009 ID T25009 standard; cDNA to mRNA; 108 BP.
AC T25009;
DT 07-NOV-1996 (first entry)
DE Human gene signature HUMGS07131.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-AL.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 1748; 2245pp; Japanese.

CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

SQ Sequence 108 BP; 34 A; 31 C; 26 G; 15 T;

Query Match 0.2%; Score 58.6; DB 1; Length 108;
Best Local Similarity 71.0%; Pred. No. 0.093;
Matches 76; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 7325 ATGCGGTGAACCGGGAGCGGAGCTTTCAGTGTAGCGGAGATCGCGCATGGCACTCCAG 7384
DB 2 ATCGCTTGGCCCATATGAGGCCAAGGCTGCTGAGTGGATGGTTCAGCTGATTCAG 61

QY 7385 CCTGGTGTACAGCGAGACTCCGCTCTCAAAAAAATAAAAAA 7431
DB 62 CCTGAGTGACAGAGCAAGACCCCTGTTGAAAAACAACAACANCAA 108

RESULT 10
T26288

ID T26288 standard; cDNA to mRNA; 110 BP.
AC T26288;
DT 10-OCT-1996 (first entry)
DE Human gene signature HUMGS08527.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 2048; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 110 BP; 22 A; 37 C; 28 G; 17 T;

Query Match 0.2%; Score 58.6; DB 1; Length 110;
Best Local Similarity 85.3%; Pred. No. 0.093;
Matches 64; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 9847 GATCCACCGCGCTCGGCTCCCAAGCTGCTGGATTACAGGTATGAGCCACTGGGCGCGG 9906
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 GATCCGCGCGCTCACCTCCCAAGCTGCTGGATTACAGGTATGAGCCACTGCACCGCG 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 9907 CCACATTTCTAAAT 9921
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 CCCCATTCCTCACTT 75
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 11
T25260/c
ID T25260 standard; DNA; 110 BP.
AC T25260;
DT 11-NOV-1996 (first entry)
DE Human gene signature HUMGS07421.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PT tissues
PS Claim 1; Page 1807; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 110 BP; 34 A; 27 C; 23 G; 20 T;

Query Match 0.2%; Score 58.4; DB 1; Length 110;
Best Local Similarity 84.4%; Pred. No. 0.1;
Matches 65; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 9607 TTTTGTGTTTGTGTTTGTGAGCGAGTTTCACACTTGTGTCAGGCTGGAGTGCAA 9666
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 77 TGTGTTGTTGTTGTTGTCGTGGAGCGAGTTGCTCTTGTGTCAGGCTGGAGTGCAA 18
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 9667 TGTGTCGATCTCGGCTC 9683
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 17 TGTGTCATCTCGGATC 1

RESULT 12
T26828
ID T26828 standard; cDNA to mRNA; 108 BP.
AC T26828;
DT 14-NOV-1996 (first entry)
DE Human gene signature HUMGS09078.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 2182; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 103 BP; 18 A; 33 C; 23 G; 28 T;

PT tissues
PS Claim 1; Page 2158; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 97 BP; 19 A; 27 C; 20 G; 28 T;

Query Match 0.2%; Score 54.8; DB 1; Length 97;
Best Local Similarity 73.9%; Pred. No. 0.41;
Matches 68; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 14573 GATCTGCCCGCCTTGGCCTCCCAAAGTGTGGGATTACAGGCATCAGCCACGTCGCCCAG 14632
DB 1 GATCTGCCCACCTNGGCCTCCCGAGTGCTGGGATTACAGGCATGAGCCACTGCCCCGG 60

QY 14633 CCAGGAGCAGATTTTATCACTCATGTTTCT 14664
DB 61 NCTGTACTAAGTCTTTTTTTTAAATTCCT 92

Search completed: June 16, 2000, 00:47:51
Job time: 128982 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2000, 08:12:30 ; Search time 8513.49 Seconds
(without alignments)
13807.208 Million cell updates/sec

Title: US-08-852-495C-1_COPY_84000_113000
Perfect score: 29001
Sequence: 1 TTGTATTTTATTAGAGACA.....GCCTCAGCCTTCAGAGTGTA 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues
Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : EST:
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result	%	Query
SUMMARIES		

Seq primer: -28ml3 rev1 ET from Amersham	
High quality sequence stop: 53.	
Location/Qualifiers	
1. .106	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone="IMAGE:1140858"	
/clone_lib="Stratagene hNT neuron (#937233)"	
/dev_stage="hNT neurons"	
/lab_host="SOLR (kanamycin resistant)"	
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Differentiated, post mitotic hNT neurons. Average insert size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"	
BASE COUNT	19 a 29 c 29 g 29 t
ORIGIN	
Query Match	0.3%; Score 94.8; DB 37; Length 106;
Best Local Similarity	93.4%; Pred. No. 0.094;
Matches	99; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2767	CACGCCGTGTAATCTAGCACTTTGGGAGCCGAGACGGCGGCATCACGAGTCAGGAGAT 2826
Db 106	CACGCCGTGTAATCCAGCACCTTGGGAGGCTGAGCGGGCAGATCACGAGTCAGGAGAT 47
QY 2827	CGAGACCATCTTGGCTAACACGGTGAACCCGTTTCTACTAAAAA 2872
Db 46	CGAGACCATCTTGGCTAACACGGTGAACCCGTCGTCCTACTAAAAA 1
RESULT 3	
AI832832	
LOCUS	AI832832 105 bp mRNA EST 13-JUL-1999
DEFINITION	at72909.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone IMAGE:2377600 3' similar to contains Alu repetitive element; contains element MER22 repetitive element ;, mRNA sequence.
ACCESSION	AI832832
VERSION	AI832832.1 GI:5454812
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 105) Hillier L., Allen M., Bowles L., Dubuque T., Geisels G., Jost S., Krizman D., Kucaba T., Lacy M., Le N., Lennon G., Marra M., Martin J., Moore B., Schellenberg K., Steptoe M., Tan F., Theising B., White Y., Wyllie T., Waterston R. and Willson R. WashU-NCI human EST Project Unpublished (1997)
TITLE	Unpublished (1997)
JOURNAL	On Dec 20, 1995 this sequence version replaced gi:1133644.
COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: -400p from Gibco.
FEATURES	
source	Location/Qualifiers 1. .105 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2377600" /clone_lib="Barstead colon HPLRB7" /sex="male" /dev_stage="adult, age 25" /lab_host="DH10B (phage resistant)" /note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with

Query Match 0.3%; Score 91.8; DB 84; Length 103;
 Best Local Similarity 93.2%; Pred. No. 0.21;
 Matches 96; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2774 GTAATCTAGCACTTTGGAGCCGAGACGGCGGATCACGAGGTCAGGATCGAGACC 2833
 ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1 GTAAGCCAGCACTTTGGAGCCGAGACGGCGGATCACGAGGTCAGGATCGAGACC 60

QY 2834 ATCTTGGCTATACACGCTGAACCCGCTTCTACTATAAATACA 2876
 ||||| ||||||||||||||||||||||||||||||||||||||||||||||||
 DB 61 ATCCCGCTAAACCGTGAACCCGCTCTCTACTATAAATACA 103

RESULT 5
 AA812141/c
 LOCUS AA812141 106 bp mRNA EST 19-FEB-1998
 DEFINITION ob48h02.s1 NCI-CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1334643 3'
 similar to contains Alu repetitive element; , mRNA sequence.
 ACCESSION AA812141
 VERSION AA812141.1 GI:2881752
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 106)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1402063.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
 Ph.D., Gerald Marti, M.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 1450 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 60.

FEATURES

source
 1..106
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1334643"
 /clone_lib="NCI-CGAP_GCB1"
 /tissue_type="germinal center B cell"
 /lab_host="DH10B"
 /note="vector: p7T3P-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from human tonsillar cells enriched for
 germinal center B cells by flow sorting (CD20+, IgD-),
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
 primed with a Not I - oligo(dT) primer
 [5'-TGTTACCAATCTGAAGTGGGCGGCGCTCATTTTTTTTTTTT-
 3']. Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 16 a 31 c 24 g 35 t
 ORIGIN
 |||||||

Query Match 0.3%; Score 91.6; DB 38; Length 106;

Best Local Similarity 91.5%; Pred. No. 0.21;
 Matches 97; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 7320 GGAGAATGGCTGAACCGGCGGAGCGGAGCTTGAGTCAGCCGAGATCGCCCATGGCAC 7379
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 DB 106 GGAGAATGGCTGAACCTGGGAGGTGGAGCTTGAGTCAGCCGAGATCACCATGGCAC 47

QY 7380 TCCAGCCTGGGTGACAGAGCGAGACTCCGTCTCAAAAAA 7425
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 DB 46 TCCAGCCTGGGTGACAGAGCGAGACTCCATCTCAAAAAA 1

RESULT 6
 AA828124
 LOCUS AA828124 107 bp mRNA EST 20-FEB-1998
 DEFINITION oD71a07.s1 NCI-CGAP_Ov2 Homo sapiens cDNA clone IMAGE:1373364
 similar to contains Alu repetitive element; contains element MER22
 repetitive element ; , mRNA sequence.
 ACCESSION AA828124
 VERSION AA828124.1 GI:2900487
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 107)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Jan 17, 1998 this sequence version replaced gi:1899815.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: David B. Krizman, Ph.D.
 DNA Library Arrayed by: Greg Lennon, Ph.D.
 CDNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 93.
 Location/Qualifiers
 1..107

FEATURES
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1373364"
 /clone_lib="NCI-CGAP_Ov2"
 /sex="female"
 /tissue_type="ovary"
 /lab_host="DH10B"
 /note="vector: pAMP10; mRNA made from invasive ovarian
 tumor, cDNA made by oligo-dT priming. Non-directionally
 cloned. Size-selected on agarose gel, average insert size
 600 bp. Reference: Krizman et al. (1996) Cancer Research
 56:5380-5383."

BASE COUNT 30 a 23 c 38 g 16 t
 ORIGIN
 |||||||

Query Match 0.3%; Score 91.2; DB 39; Length 107;
 Best Local Similarity 92.3%; Pred. No. 0.24;
 Matches 96; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2862 TCTACTAAAAATACAAAAAATAGCCGGCGTGTGGGGCGGCGCTGTAGTCCACGCTAC 2921
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 DB 4 TCGACTAAAAATACAAAAAATAGCCAGCGTAAATGGCGGCACCTGTAGTCCAGCTGC 63

QY 2922 TTGGAGGCTGAGGAGGAGATGGCATGACCTGGGAGGGGA 2965
 |||||||


```
Db 64 TTGGGAGGCTGAGCAGGAGATGGCGTGAACCCGGAGCGGA 107
RESULT 7
AA243009
LOCUS AA243009 109 bp mRNA EST 11-MAR-1998
DEFINITION cDNA clone IMAGE:664467 3' similar to contains Alu repetitive
element;contains element LTR1 repetitive element ;, mRNA sequence.
ACCESSION AA243009
VERSION AA243009.1 GI:1873869
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 109)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Dec 3, 1996 this sequence version replaced gi:1126869.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1127 Std Error: 0.00
Seq Primer: 41m13 fwd. ET from Amersham
High quality sequence stop: 102.
FEATURES
Location/Qualifiers
1..109
/organism="Homo sapiens"
/db_xref="GDB:5426481"
/db_xref="taxon:9606"
/clone="IMAGE:664467"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/tissue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/notes="Organ: brain; Vector: pluescript SK-; Site_1:
EcoRI; Site_2: xhoI; Cloned unidirectionally. Primer:
Oligo dt. Uninduced, exponentially growing neuroepithelial
cells (Ntera-2/cl.D1). Average insert size: 1.0 kb;
Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCAGG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3"
```

```
BASE COUNT 19 a 30 c 30 g 30 t
ORIGIN
Query Match 0.3%; Score 91.4; DB 30; Length 109;
Best Local Similarity 89.9%; Pred. No. 0.22;
Matches 98; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 8700 GTATTTTATGACAGAGGGTTTACCGTGTGGCCAGAGTGTCTCAATCTCCTTACC 8759
Db 1 GTATTTTATGACAGAGGGTTTACCGTGTGGCCAGAGTGTCTTGTGATCTCCCTACC 60

QY 8760 TCGTGATCGCCCGCTCTCTCTGCAAAAGTCTCGGATTACAGCGTG 8808
Db 61 TCGTGATCGCCCGCTCTCTCTGCAAAAGTCTCGGATTACAGCGGTG 109

RESULT 8
AA835205/c
LOCUS AA835205 101 bp mRNA EST 23-FEB-1998
DEFINITION ak64h01.s1 Barstead pancreas HPLRB1 Homo sapiens cDNA clone
IMAGE:1412689 3' similar to contains Alu repetitive
element;contains element KER repetitive element ;, mRNA sequence.
ACCESSION AA835205
VERSION AA835205.1 GI:2908933
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 101)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Nov 29, 1993 this sequence version replaced gi:636191.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq Primer: -40m13 fwd. ET from Amersham.
FEATURES
Location/Qualifiers
1..101
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1412689"
/clone_lib="Barstead pancreas HPLRB1"
/sex="female"
/dev_stage="adult, 34 years"
/lab_host="DH10B"
/notes="Organ: pancreas; Vector: pTT73D-Pac (Pharmacla)
with a modified polylinker; Site_1: EcoRI; Site_2: NotI;
1st strand cDNA was primed with a Not I - olligo(dT) primer
[5',
TGTTACGAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[AAATCGGATCCTTG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pTT73 vector.
Library constructed by Bob Barstead."
BASE COUNT 14 a 36 c 27 g 24 t
ORIGIN
Query Match 0.3%; Score 89.8; DB 39; Length 101;
Best Local Similarity 93.1%; Pred. No. 0.35;
Matches 94; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 7313 TGAGGCAGAGATGCGGTGAACCGGGAGCGGAGCTTGCAGTGCAGATCGCGCC 7372
Db 101 TGAGGCAGAGATGCGGTGAACCGGGAGCGGAGCTTGCAGTGCAGATCAAGCC 42

QY 7373 ATGGCACTCCAGCTGGGTGACAGAGCGAGACTCCGCTCTCA 7413
Db 41 ACTGCACTCCAGCTGGGCGACAGAGTGCAGATCCGCTCTCA 1

RESULT 9
N25299/c
LOCUS N25299 109 bp mRNA EST 28-DEC-1995
DEFINITION yw52c09.s1 Weizmann Olfactory Epithelium Homo sapiens cDNA clone
IMAGE:255856 3' similar to contains Alu repetitive element; , mRNA
sequence.
ACCESSION N25299
VERSION N25299.1 GI:1139449
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```

REFERENCE
AUTHORS
1 (bases 1 to 109)
Hillier, L., Lennon, G., Becker, M., Bonaudo, M.F., Chiapelli, B.,
Chissole, S., Dietrich, N., DuBuque, T., Favella, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Marra, M.

TITLE
JOURNAL
MEDLINE
COMMENT
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
9704478
On Apr 14, 1993 this sequence version replaced gi:8377394.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: m13 -40 forward
High quality sequence stop: 307.

FEATURES
source
1..109
/organism="Homo sapiens"
/db_xref="GDB:3866265"
/db_xref="taxon:9606"
/clone_lib="IMAGE:255856"
/clone_lib="Weizmann Olfactory Epithelium"
/sex="Female"
/tissue_type="olfactory epithelium"
/lab_host="35 year old"
/lab_host="SOLR cells (kanamycin resistant)"
/notes="Organ: nose; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Olfactory epithelium, normal. Average insert
size: 0.8 kb; Uni-ZAP XR Vector. Library constructed by N.
Walker, D. Lancet, Weizmann Institute of Science. -5'
adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"
BASE COUNT      13 a      34 c      24 g      35 t      3 others
ORIGIN

Query Match      0.3%; Score 90; DB 24; Length 109;
Best Local Similarity 88.1%; Pred. No. 0.32;
Matches 96; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 2933 AGGCAGGAGATGGCATGAACCTGGGAGCGAGCTTGCAGTGCAGCGGAGATTGGCCAC 2992
|||||
Db 109 AGGCAGGAGATGGCAGAACCTGGGAGCGAGCTTGCAGTGCAGCGGAGATTGGCCAC 50
|||||

Qy 2993 TGCACCTCCAACTGGGAGACACAGCGAGACTCCGCTCAAAAAA 3041
|||||
Db 49 TGCACCTCCAGCTGGGAGACAGAGCGAAGTCCGCTCAAAAAA 1
|||||

RESULT 10
N49638
LOCUS
DEFINITION
y25e09.r1 Soares fetal liver spleen lNFLS Homo sapiens cDNA clone
IMAGE:243784 5' similar to gb:X57138_rnal HISTONE H2B.2 (HUMAN);,
mRNA sequence.
ACCESSION
N49638
VERSION
N49638.1 GI:1190804
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 97)
AUTHORS
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,

Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
On Apr 14, 1993 this sequence version replaced gi:693230.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Seq primer: T7
High quality sequence stop: 1.

FEATURES
source
1..97
/organism="Homo sapiens"
/db_xref="GDB:3792917"
/db_xref="taxon:9606"
/clone_lib="IMAGE:243784"
/clone_lib="Soares fetal liver spleen lNFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAAGAATAATTAAGATCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      24 a      29 c      23 g      20 t      1 others
ORIGIN

Query Match      0.3%; Score 89.6; DB 25; Length 97;
Best Local Similarity 94.8%; Pred. No. 0.37;
Matches 92; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 6698 GTACAAAGTTCTGAGCAGGTCCACCCGACACCGGCATCTCATCCAGGCGCATGGGAT 6757
|||||
Db 1 GTACAAAGTTCTGAGCAGGTCCATCCGACACCGGCATCTCTCCAAAGGCAATGGGAT 60
|||||

Qy 6758 CATGAATTCCTTCGTCAACGACATCTTCGAGCGCATC 6794
|||||
Db 61 CATGAATTCCTTCGTCAACGACATCTTCGAGCGCATC 97
|||||

RESULT 11
AQ028649
LOCUS
DEFINITION
CIT-HSP-2323P12.TR CIT-HSP Homo sapiens genomic clone 2323P12,
genomic survey sequence.
ACCESSION
AQ028649
VERSION
AQ028649.1 GI:3268871
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 103)
AUTHORS
Adams, M.D., Rounsley, S.D., Zhao, S., Field, C.E., Bass, S., Linher, K.,
Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H.,
Simon, M. and Venter, J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
Unpublished (1998)
JOURNAL

```

Other_GSSs: CIT-HSP-2323P12.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES	source	location/Qualifiers
BASE COUNT	35 a	27 c 28 g 13 t
ORIGIN		
		1.103
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="2323P12"
		/clone_lib="CIT-HSP"
		/sex="Male"
		/cell_type="Sperm"
		/note="vector: pBeloBAC11; Site_1: HindIII; Site_2: HindIII"

[illegible]

RESULT	12
AQ076649/c	
LOCUS	AQ076649 101 bp DNA GSS 20-AUG-1998
DEFINITION	CIT-HSP-2363C23.TR CIT-HSP Homo sapiens genomic clone 2363C23, genomic survey sequence.
ACCESSION	AQ076649
VERSION	AQ076649
KEYWORDS	AQ076649.1 GI:3437833
SOURCE	GSS.
ORGANISM	human. Homo sapiens Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 101)
AUTHORS	Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.
TITLE	Use of a random human BAC End Sequence Database for Sequence-Ready

JOURNAL
COMMENT

Unpublished (1998)
Other_GSSs: CIT-HSP-2363C23.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES

Location/Qualifiers

```

source
1. .101
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="J363C23"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/notes=Vector: pBelOAc11; Site_1: HindII; Site_2:
HindIII"
12 a 29 c 28 g 32 t
BASE COUNT
ORIGIN

```

	Query Match	0.3%	Score 88.2	DB 94	Length 101
	Best Local Similarity	92.1%	Pred. No. 0.52		
	Matches 93	Conservative 0	Mismatches 8	Indels 0	Gaps 0
Qy 7323	GAATGGCGTGAA	CCGGAGCGGAGCTTG	CAGTGTGAGCCGAGATCGG	CCCATCGCACTCC	7382
Db 101	GAATGGCGTGAA	CCCGAGCGGAGCTTG	CAGTGTGAGCCGAGATCGG	CCCATCGCACTCC	42
Qy 7383	AGCCTGGGTGAC	AGAGCGGAGCTCCG	TCTCAAAAAAAAAA	7423	
Db 41	AGCCTGGGTGAC	AGAGCGGAGCTCCG	TCTCAAAAAAAAAA	1	

RESULT	13
LOCUS	AQ264176
LENGTH	106 bp
DNA	GSS
DATE	27-OCT-1998
DEFINITION	CITBI-EI-2509A2.TF CITBI-EI Homo sapiens genomic clone 2509A2, genomic survey sequence.
ACCESSION	AQ264176
VERSION	AQ264176.1
KEYWORDS	GI:3792743
SOURCE	GSS. human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 106)
AUTHORS	Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.

TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
JOURNAL Map Building
COMMENT Unpublished (1998)
 Other_GSSs: CITBI-E1-2509A2.TR
 Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: M13-21
 Class: BAC ends.

FEATURES	Class. and ends.	Location/Qualifiers
source	1. .106	/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="2509A2"
		/clone_lib="CIBI-EI"
		/sex="male"
		/cell_type="sperm"
		/notes=Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
BASE COUNT	25 a 30 c 34 g 17 t	Caltech Human BAC library D"
ORIGIN		
Query Match	0.3%	Score 88.4; DB 105; Length 106;

Best Local Similarity 89.6%; Pred. No. 0.48;
Matches 95; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2752 CCGGCGCGGTGGCTACGCTGTAACTTACGACTTTGGGAGCGGAGCGGGGATC 2811

Db 1 CCGGCGCGAGAGTCACGCTGTAACTTACGACTTTGGGAGCGGAGCGGGGATC 60

Qy 2812 ACGAGTCAGGAGTCGAGACCATCTTGGCTAACACGCTGAAACCC 2857

Db 61 ACGAGTCAGGAGTCAGACCGCTCTGGCTAACATGCTGAAACCC 106

RESULT 14

AA807640/c

LOCUS

DEFINITION nx08b05.s1 NCI_CGAP_GC3 Homo sapiens cDNA clone IMAGE:1255473 3'

similar to contains Alu repetitive element;; mRNA sequence.

ACCESSION AA807640

VERSION AA807640.1 GI:2877108

KEYWORDS EST.

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 103)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL Unpublished (1987)

COMMENT On Jan 19, 1998 this sequence version replaced gi:2151346.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 774 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 87.

FEATURES

source

1. .103

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1255473"

/clone_lib="NCI_CGAP_GC3"

/tissue_type="pooled germ cell tumors"

/lab_host="DH10B"

/note="Vector: p7T3D-Pac (Pharmacia) with a modified

polylinker; 1st strand cDNA was prepared from 3 pooled

germ cell tumors, and was then primed with a Not I -

oligo(dT) primer. Double-stranded cDNA was ligated to Eco

RI adaptors (Pharmacia), digested with Not I and cloned

into the Not I and Eco RI sites of the modified p7T3

vector. Library is not normalized. Library was

constructed by Bento Soares and M. Fatima Bonaudo. "

19 a 27 c 30 g 27 t

BASE COUNT

ORIGIN

Query Match 0.3%; Score 87.6; DB 38; Length 103;

Best Local Similarity 91.2%; Pred. No. 0.6;

Matches 93; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 2766 TCACGCTGTAACTTGGGAGCGGAGCGGATCAGGTCAGGAGA 2825

Db 103 TCACGCTGTAACTTGGGAGCGGAGCGGATCAGGTCAGGAGA 44

Qy 2826 TCGAGACCATCTTGGCTAACACGCTGAAACCCGCTTCTACT 2867

Db 43 TCGAGACCATCTTGGCTAACACGCTGAAACCCGCTTCTACT 2

RESULT 15

AW250394

LOCUS

DEFINITION 2822460.3prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2822460 3',

mRNA sequence.

ACCESSION AW250394

VERSION AW250394.1 GI:6593387

KEYWORDS EST.

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 110)

NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT On May 18, 1998 this sequence version replaced gi:3138342.

Other ESTs: 2822460.5prime

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: DCTD/DTF CDNA Library Preparation: Ling

Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E.

Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing

project Clone distribution: MGC clone distribution information can

be found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html Base Calling / Quality

Scores: PHRED high quality bases following vector sequence. Very

contiguous PHRED high quality bases following vector sequence. Very

Low Quality Sequence: Trace file contained 110 contiguous distinct

peaks following vector sequence. Polyadenylation: Based upon the

presence of a XhoI site followed by a run of 14 or more T residues

at the beginning of the sequence, this cDNA insert was

polyadenylated.

Plate: LLCM9 row: H column: 13

High quality sequence stop: 61.

FEATURES

source

Location/Qualifiers

1. .110

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2822460"

/clone_lib="NIH_MGC_7"

/tissue_type="small cell carcinoma"

/cell_line="MGC3"

/lab_host="DH10B (phage-resistant)"

/note="Organ: lung; Vector: pOTB7; Site:1: XhoI; Site:2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

19 a 29 c 24 g 38 t

BASE COUNT

ORIGIN

Query Match 0.3%; Score 87.6; DB 79; Length 110;

Best Local Similarity 87.3%; Pred. No. 0.58;

Matches 96; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 9617 TTTTTCGAGCGGAGTTTCACACTTTGCCAGCTGGAGTGCATGTCGATC 9676

Db 1 TTTTTCGAGCGGAGTTTCGCTCTATTGCCCCAAGCTGGAGTGCATGTCAT 60

Oy 9677 TCGGCTACCGCAACTCGGCCTCCAGGTTCAAGCAATTCCTCGCCTC 9726
||||| ||| | ||||| ||||| || ||||| ||||| ||||| ||||| |||||
Db 61 TCGGCTCAACTCAACCTCGGCCTCCGGTTTCAAGAGATTCTCTCGCCTC 110

Search completed: June 15, 2000, 19:11:36
Job time: 110063 sec


```

CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,862B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-485-862B-65

Query Match 0.3%; Score 80.2; DB 4; Length
Best Local Similarity 87.1%; Pred. No. 2.5e-09;
Matches 88; Conservative 0; Mismatches 13; Indel

QY 1 TTGATTTTATTAGACAGGGTTTCACTATGTTGCCAGGCTGATCTC
DB 5 TTACTCTTTAGTAGACAGGGTTTTCACCATATTGCCAGGCTGCTC
QY 61 CPTCATGATCCGCTCGCTTGGCTCTCAAAGTCTGGGAT 101
DB 65 CCTTGATGATCCAGGCTCGGCTCCCAAAGTCTGGGAT 105

RESULT 5
US-08-787-739-65
Sequence 65, Application US/08787739
Patent No. 6027887
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN gene and Protein
NUMBER OF SEQUENCES: 96
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 369 Pine Street, Suite 610
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-787-739-65

Query Match 0.3%; Score 80.2; DB 5; Length 105;
Best Local Similarity 87.1%; Pred. No. 2.5e-09;
Matches 88; Conservative 0; Mismatches 13; Indels 0; Gaps

Qy 1 TTCTATTATTATAGACAGGGTTTCACTATGTTGCCAGGCTGATCTCAAACTCCTGA 60
Db 5 TTACTTTTATTAGACAGGGTTTTCACCATATTTGGCCAGGCTGCTCTCAAACTCCTGA 64
Qy 61 CCTCATGATCCGCCCTCGCCTCTCAAAAGTCTCGGAT 101
Db 65 CCTGTGATCCACAGCCTCGGCCCTCCCAAAGTCTGGGAT 105

RESULT 6
US-08-481-6588-65/c
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.3%; Score 76.2; DB 4; Length 105;
Best Local Similarity 82.9%; Pred. No. 2e-08;
Matches 87; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 2777 ATCTAGCACTTTGGGAGCGGCGGATCAGAGTCAAGGATCGAGGATCGAGACCATC 2836
|||||
Db 105 ATCCAGCACTTTGGGAGCGGCGGATCAGAGTCAAGGATCGAGGATTTGAGAGCAGC 46
|||||

Qy 2837 TTGGCTAACACGGTGAACCCGTTTCTACTAAAAATACAAAAA 2881
|||||
Db 45 CTGCCAATATGGTGAACCCCTGCTCTACTAAAGATGTAAAAA 1

RESULT 7
US-08-477-504A-65/c
Sequence 65, Application US/08477504A
Patent No. 5972353
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-477-504A-65

Query Match 0.3%; Score 76.2; DB 4; Length 105;
Best Local Similarity 82.9%; Pred. No. 2e-08;
Matches 87; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 2777 ATCTAGCACTTTGGGAGCGGCGGATCAGAGTCAAGGATCGAGGATCGAGACCATC 2836
|||||
Db 105 ATCCAGCACTTTGGGAGCGGCGGATCAGAGTCAAGGATCGAGGATTTGAGAGCAGC 46
|||||

Qy 2837 TTGGCTAACACGGTGAACCCGTTTCTACTAAAAATACAAAAA 2881
|||||
Db 45 CTGCCAATATGGTGAACCCCTGCTCTACTAAAGATGTAAAAA 1

RESULT 8
US-08-486-756A-65/c
Sequence 65, Application US/08486756A
Patent No. 5981711
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,756A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear


```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450.673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
; US-08-450-673C-91
;
; Query Match 0.2%; Score 63.8; DB 4; Length 84;
; Best Local Similarity 85.5%; Pred. NO. 1.1e-05;
; Matches 71; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
;
; Qy 2767 CACGCCTGTAATCCTAGCACACCTTTGGGAGCGCGGATCACAGGTCAGAGAT 2826
; Db 83 CACGCTTGTATCCAGCACACCTTTGGGAGCGCTGAGCGCGGATCACAGGTCAGGAGTT 24
;
; Qy 2827 CGAGACCATCTTGGCTAACACGG 2849
; Db 23 CGACACCAGCCTGATGAACATGG 1
;
; RESULT 14
; PCT-US95-17111A-91/c
; SEQUENCE 91, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
; US-08-450-673C-91
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340.426
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
```

```
;
; TOPOLOGY: both
; PCT-US95-17111A-91
;
; Query Match 0.2%; Score 63.8; DB 6; Length 84;
; Best Local Similarity 85.5%; Pred. NO. 1.1e-05;
; Matches 71; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
;
; Qy 2767 CACGCCTGTAATCCTAGCACACCTTTGGGAGCGCGGATCACAGGTCAGAGAT 2826
; Db 83 CACGCTTGTATCCAGCACACCTTTGGGAGCGCTGAGCGCGGATCACAGGTCAGGAGTT 24
;
; Qy 2827 CGAGACCATCTTGGCTAACACGG 2849
; Db 23 CGACACCAGCCTGATGAACATGG 1
;
; RESULT 15
; US-08-454-557C-69
; SEQUENCE 69, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454.557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
; US-08-454-557C-69
;
; Query Match 0.2%; Score 62.8; DB 3; Length 76;
; Best Local Similarity 97.0%; Pred. NO. 1.8e-05;
; Matches 64; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; Qy 8680 ACCATGCCCGCTAAATTTTGTATTTTAGTAGACACAGGGTTTACCCTGTGGCCAGG 8739
; Db 1 ACCACGCCCGCTAAATTTTGTATTTTAGTAGACACAGGGTTTACCCTGTGGCCAGG 60
;
; Qy 8740 ATGCTC 8745
; Db 61 ATGCTC 66
```

Search completed: June 16, 2000, 00:17:42
Job time: 127392 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2000, 00:11:20 ; Search time 17971 Seconds
(without alignments)
-1569.858 Million cell updates/sec

Title: US-08-852-495C-1_COPY_112000_141000
Perfect score: 29001
Sequence: 1 TTTTCCACRCTTTCTTCAG.....AGAGTGTTCACCTCTAGGA 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba1.*
- 2: gb_ba2.*
- 3: gb_om.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_pl1.*
- 8: gb_pl2.*
- 9: gb_pr1.*
- 10: gb_pr2.*
- 11: gb_pr3.*
- 12: gb_ro.*
- 13: gb_sts.*
- 14: gb_sy.*
- 15: gb_un.*
- 16: gb_v1.*
- 17: gb_fun.*
- 18: gb_hum1.*
- 19: gb_hum2.*
- 20: gb_in.*
- 21: gb_om.*
- 22: gb_or.*
- 23: gb_ov.*
- 24: gb_pat.*
- 25: gb_ph.*
- 26: gb_pl.*
- 27: gb_ro.*
- 28: gb_sts.*
- 29: gb_sy.*
- 30: gb_un.*
- 31: gb_v1.*
- 32: gb_htg1.*
- 33: gb_htg2.*
- 34: gb_in1.*
- 35: gb_in2.*
- 36: gb_ba1.*
- 37: gb_ba2.*
- 38: gb_hum3.*
- 39: gb_hum4.*
- 40: gb_pr4.*
- 41: gb_htg3.*
- 42: gb_htg4.*
- 43: gb_htg5.*
- 44: gb_htg6.*

- 45: gb_htg7.*
- 46: em_htg1.*
- 47: em_htg2.*
- 48: em_htg3.*
- 49: em_hum5.*
- 50: gb_pl3.*
- 51: gb_pr5.*
- 52: gb_htg8.*
- 53: gb_htg9.*
- 54: gb_htg10.*
- 55: gb_htg11.*
- 56: gb_htg12.*
- 57: gb_htg13.*
- 58: gb_htg14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	99	0.3	108	10	HSLDLRN2	X05250 Human LDL-r
2	90.2	0.3	107	9	HUMALCE162	M87924 Human carci
3	89.4	0.3	108	10	HSLDLRD1	X05249 Human LDL-r
4	89.4	0.3	108	10	HSLDLRD2	X05251 Human LDL-r
5	86.8	0.3	108	10	HSLDLRN2	X05250 Human LDL-r
6	85.4	0.3	103	9	HUMALCE221	M87896 Human carci
7	83.2	0.3	108	10	HSLDLI12	X05248 Human LDL-r
8	82.6	0.3	108	11	HSU67803	U67803 Human small
9	81.2	0.3	107	9	HUMALCE162	M87924 Human carci
10	79.8	0.3	103	9	HUMALCE221	M87896 Human carci
11	79.4	0.3	108	10	HSLDLRD1	X05249 Human LDL-r
12	79.4	0.3	108	10	HSLDLRD2	X05251 Human LDL-r
13	76	0.3	103	13	HS8IC8R	X57789 Human sequ
14	75.6	0.3	110	11	HSU67807	U67807 Human small
15	75.2	0.3	103	13	HS8IC8R	X57789 Human sequ
16	74.4	0.3	104	9	HUMALCE272	M87899 Human carci
17	74.2	0.3	108	9	HUMDL003M5	D16965 Human Hepo2
18	74.4	0.3	108	11	HSU67808	U67808 Human small
19	73.6	0.3	97	9	HUMDLRA2	M14180 Human low d
20	73.4	0.3	110	11	HSU67807	U67807 Human small
21	71.4	0.2	108	11	HSU67804	U67804 Human small
22	70.8	0.2	99	13	HUMUT7692A	L30306 Human STS U
23	70.8	0.2	108	10	HSLDLI12	X05248 Human LDL-r
24	70	0.2	95	13	HOMUT8002B	L30176 Human STS U
25	70.2	0.2	108	13	G32614	G32614 A009K21 Hum
26	69.8	0.2	101	10	S79560	S79560 HRX (intrn
27	69.6	0.2	107	11	HSU67806	U67806 Human small
28	69.4	0.2	108	13	G43535	G43535 WIAF-2393-S
29	69	0.2	90	9	HUMDLRLFL	K03555 Human low d
30	69.2	0.2	91	13	HOMUT8164A	L30244 Human STS U
31	68.4	0.2	79	10	S73203	S73203 ALL-1 (tand
32	68.2	0.2	101	10	S79560	S79560 HRX (intrn
33	67.8	0.2	95	10	HSSTHPK1B	X66361 H.sapiens m
34	68	0.2	108	13	G43535	G43535 WIAF-2393-S
35	67.2	0.2	80	9	HUMBRKFAE	M36135 Human alpha
36	66.6	0.2	91	13	HOMUT8164A	L30244 Human STS U
37	66.8	0.2	94	9	HUMHGAL	M13479 Human alpha
38	66.2	0.2	100	9	HUMGALNSA	D45223 Human GALNS
39	66.2	0.2	100	9	HUMGALNSA	D45223 Human GALNS
40	65.6	0.2	97	9	HUMDLRA2	M14180 Human low d
41	65.2	0.2	108	9	HUMDL003M5	D16965 Human Hepo2
42	64.4	0.2	90	10	HSU19407	U19407 Human isolat
43	64.6	0.2	100	10	HSLAS27	X91545 H.sapiens D
44	64.4	0.2	110	9	HUMALCE43	M87900 Human carci
45	64	0.2	90	9	HUMDLRLFL	K03555 Human low d

ALIGNMENTS

```

RESULT 1
HSLDLR2  HSLDLR2  108 bp  DNA  PRI  20-MAY-1992
LOCUS    Human LDL-receptor gene intron 14 fragment (normal gene).
DEFINITION
ACCESSION X05250
VERSION   X05250.1 GI:34337
KEYWORDS  Alu repetitive sequence; low density lipoprotein receptor.
SOURCE    human.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
           Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
           Williamson,R., and Humphries,S.
TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
           the low-density-lipoprotein-receptor gene. A possible mechanism for
           the defect in a patient with familial hypercholesterolaemia
JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE   87161901
COMMENT   See X05252 for deletion junction
           Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES  source
           location/Qualifiers
           1..108
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           1..108
           /note="intron XIV fragment"
BASE COUNT 28 a 23 c 39 g 18 t
ORIGIN
Query Match 0.3%; Score 99; DB 10; Length 108;
Best Local Similarity 95.3%; Pred. No. 1e-07; Mismatches 0; Gaps 0;
Matches 102; Conservative 0; Indels 5; Indels 0; Gaps 0;

Qy 25347 ACAAAATCAGCAGCGGTGGTGGCATGTGCTGTAAATCCAGCTACTCAGGAGCTGAG 25406
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 ACAAAATAGCAGCGGTGGTGGCATGTGCTGTAAATCCAGCTACTCAGGAGCTGAG 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 25407 GCAAGAGATTCCTTGACCCAGGAGCGGAGGTTCAGTGCAGTGCAGCGCA 25453
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 GCAGGAGAATTCCTTGACCCAGGAGCGGAGGTTCAGTGCAGTGCAGCGCA 107
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 2
HUMALCE162/c HUMALCE162 107 bp ss-RNA PRI 15-APR-1994
LOCUS    Human carcinoma cell-derived Alu RNA transcript, clone CE162.
DEFINITION
ACCESSION M87924
VERSION   M87924.1 GI:174871
KEYWORDS  Alu repeat.
SOURCE    Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 107)
AUTHORS   Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE     Alu RNA transcripts in human embryonal carcinoma cells. Model of
           post-transcriptional selection of master sequences
JOURNAL   J. Mol. Biol. (1992) In press
MEDLINE
FEATURES  source
           location/Qualifiers
           1..107
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           /cell_line="NTER2D1"
           /dev_stage="embryo"
           /sex="male"
           /tissue_type="carcinoma"
BASE COUNT 28 a 30 c 35 g 14 t
ORIGIN

```

```

Query Match 0.3%; Score 90.2; DB 9; Length 107;
Best Local Similarity 92.2%; Pred. No. 3.4e-06;
Matches 95; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 20076 TTTTCTGAGACGGAGTCTGTCTGTCCGCCAGGCTGGAATGCAGTGCACAAATCTCGG 20135
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 107 TTTTCTGAGACGGAGTCTGTCTGTCCGCCAGGCTGGAATGCAGTGCAGTGCAGTCTCGG 48
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 20136 TCACTGCACACCTCCGCTCCCGGATTTCAGGCCATCTCCTGCC 20178
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 47 TCACTGCACAGTCCGCTCCCGGTTTACGCCCAATCTCTCTGCC 5

RESULT 3
HSLDLR1/c HSLDLR1 108 bp DNA PRI 20-MAY-1992
LOCUS    Human LDL-receptor mutated gene with intron 12 deletion junction.
DEFINITION
ACCESSION X05249
VERSION   X05249.1 GI:34335
KEYWORDS  Alu repetitive sequence; low density lipoprotein receptor.
SOURCE    human.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
           Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
           Williamson,R., and Humphries,S.
TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
           the low-density-lipoprotein-receptor gene. A possible mechanism for
           the defect in a patient with familial hypercholesterolaemia
JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE   87161901
COMMENT   *source: hypercholesterol aemia
           See X05248 for corresponding normal gene sequence
           alu-repetitive sequences, that are in the same direction, the
           deletion eliminates exons 13 and 14 and changes the reading frame
           of the resulting spliced mRNA.
           Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES  source
           location/Qualifiers
           1..108
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           /cell_type="blood leukocytes from a patient with familial"
           misc_feature 1..108
           /note="deletion junction region intron 12/ intron 15"
BASE COUNT 20 a 40 c 20 g 28 t
ORIGIN

Query Match 0.3%; Score 89.4; DB 10; Length 108;
Best Local Similarity 89.7%; Pred. No. 4.6e-06;
Matches 96; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 25347 ACAAAATCAGCAGCGGTGGTGGCATGTGCTGTAAATCCAGCTACTCAGGAGCTGAG 25406
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 108 ACAAAATTAGCAGCGGTGGTGGCATGTGCTGTAAATCCAGCTACTCAGGAGCTGAG 49
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 25407 GCAAGAGATTCCTTGACCCAGGAGCGGAGGTTCAGTGCAGTGCAGCGCA 25453
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 48 GCAGGAAAATGGTTGAACCCAGGAGCGAGAGGTTCGTGTGAGCGCA 2

RESULT 4
HSLDLR2  HSLDLR2  108 bp  DNA  PRI  20-MAY-1992
LOCUS    Human LDL-receptor mutated gene with intron 14 deletion junction.
DEFINITION
ACCESSION X05251
VERSION   X05251.1 GI:34336
KEYWORDS  Alu repetitive sequence; low density lipoprotein receptor.
SOURCE    human.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

```


COMMENT see X05249 for deletion junction
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.

```
FEATURES
  source      1..108
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              complement(<1..65)
              /note="Alu repeat"
  intron      1..108
              /note="intron XII fragment"
  BASE COUNT 21 a 38 c 20 g 29 t
  ORIGIN
      21 a 38 c 20 g 29 t

Query Match      0.3%; Score 83.2; DB 10; Length 108;
Best Local Similarity 87.5%; Pred. No. 5.4e-05;
Matches 91; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 25350 AAAATCAGCCAGCGGTGGTGGCATGTGCTGTAAATCCAGCTACTCAGGAGCTGAGGCA 25409
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 105 AAAATCAGCCGGCGGTGGTGGCACATCTGTGTAAATCCAGCTACTAAGGAGCTGAGGCA 46
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 25410 AGAGAATTGCTTGAACCCAGGAGGCGGAGGTTGCGAGTGAGCGGA 25453
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 45 GGAATAATGTTTGAACCCAGGAGGAGGTTGTGGTGAGGCGA 2

RESULT 8
HSU67803/c 108 bp RNA PRI 01-AUG-1997
LOCUS HSU67803 Human small cytoplasmic Alu transcript.
ACCESSION U67803
VERSION U67803.1 GI:2289917
KEYWORDS Alu.
SOURCE human.
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
    Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 (bases 1 to 108)
  AUTHORS Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE CDNA's derived from primary and small cytoplasmic Alu (scAlu) transcripts
  JOURNAL J. Mol. Biol. 271 (2), 222-234 (1997)
  MEDLINE 97415756
  REFERENCE
    2 (bases 1 to 108)
  AUTHORS Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE Direct Submission
  JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The Children's Hospital of Philadelphia, 1004F Abramson Research Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

FEATURES
  source      1..108
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="TscAlu2"
  repeat_region 1..108
              /note="scAlu"
              /rpt_family="Alu"
              /rpt_type="dispersed"
  BASE COUNT 23 a 39 c 30 g 16 t
  ORIGIN
      23 a 39 c 30 g 16 t

Query Match      0.3%; Score 82.6; DB 11; Length 108;
Best Local Similarity 90.7%; Pred. No. 6.9e-05;
Matches 88; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 20246 GTAGACAGGGTTTACCGTGTACCGGGGATGTCGATCTCCTGACCTCATGATCT 20305
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 97 GTAGACAGGGTTTACCTTCTTAGCCAGGATGTCGATCTCCTGACCTCGTGATCC 38
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 20306 GCCCACCCTCAGCCTCCCAAGTGCTAGGATCACAGC 20342
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 37 GCCGCGCTCGGCCTCCCAAGTGCTGGGATTACAGC 1
RESULT 9
HUMALCE162 107 bp ss-RNA PRI 15-APR-1994
LOCUS HUMALCE162 Human carcinoma cell-derived Alu RNA transcript, clone CE162.
ACCESSION M87924
VERSION M87924.1 GI:174871
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
    Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 (bases 1 to 107)
  AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
  TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
  JOURNAL J. Mol. Biol. (1992) In press
  FEATURES
    Location/Qualifiers
      1..107
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /cell_line="NTERa2D1"
        /dev_stage="embryo"
        /sex="male"
        /tissue_type="carcinoma"
  BASE COUNT 28 a 30 c 35 g 14 t
  ORIGIN
      28 a 30 c 35 g 14 t

Query Match      0.3%; Score 81.2; DB 9; Length 107;
Best Local Similarity 87.3%; Pred. No. 0.00012;
Matches 89; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

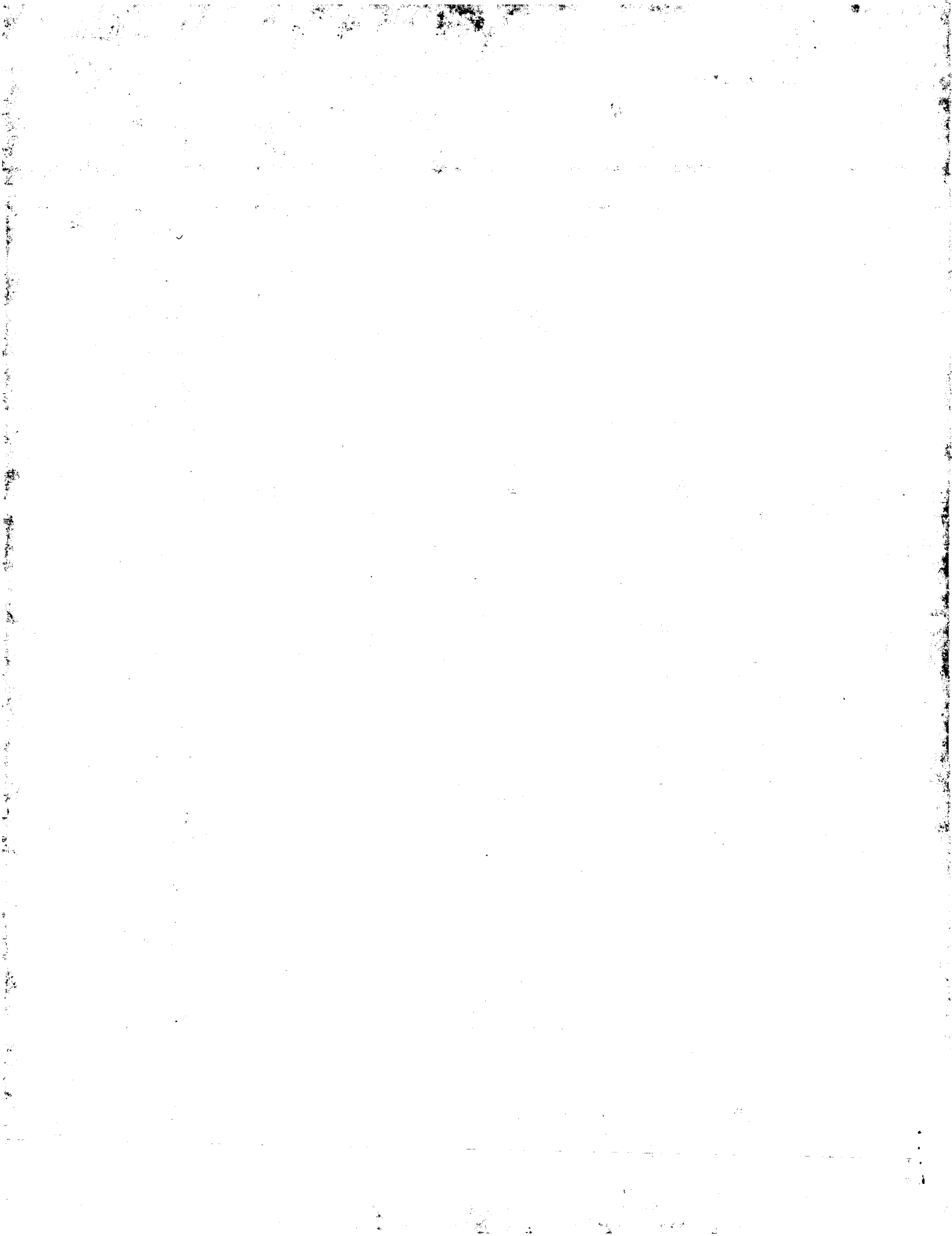
Qy 25406 GCGAAGAAATGCTTGAACCCAGGAGCGGAGGTTGCGAGTGCAGAAATCGCGCCACT 25465
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5 GCGAAGAAATGCTTGAACCCAGGAGCGGAGGTTGCGAGTGCAGAAATCGCGCCACT 64
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 25466 GCATCCAGCTGGTGAACAGACGCAAGCTCTGTTTCAAAAA 25507
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65 GCATCCAGCTGGCGGAGAGCGAGACTCCGTCCTCAAAA 106
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
HUMALCE221/c 103 bp ss-RNA PRI 15-APR-1994
LOCUS HUMALCE221 Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
    Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 (bases 1 to 103)
  AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
  TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
  JOURNAL J. Mol. Biol. (1992) In press
  FEATURES
    Location/Qualifiers
      1..103
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /cell_line="NTERa2D1"
        /dev_stage="embryo"
        /sex="male"
        /tissue_type="carcinoma"
  BASE COUNT 25 a 27 c 33 g 18 t
  ORIGIN
      25 a 27 c 33 g 18 t

Query Match      0.3%; Score 79.8; DB 9; Length 103;
```

REFERENCE	1 (bases 1 to 108)						
AUTHORS	Horsthemke,B., Betsiegel,U., Dunning,A., Haviga,J.R., Williamson,R. and Humphries,S.						
TITLE	Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia						
JOURNAL MEDLINE	Eur. J. Biochem. 164 (1), 77-81 (1987) 87161901						
COMMENT	*source: hypercholesterol aemia See X05250 for corresponding normal gene sequence In the defective LDL-receptor gene the deletion occurred between two alu-repetitive sequences, that are in the same direction, the deletion eliminates exons 13 and 14 and changes the reading frame of the resulting spliced mRNA. Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.						
FEATURES							
source	1..108 /organism="Homo sapiens" /db_xref="taxon:9606"						
intron	1..108 /cell_type="blood leukocytes from a patient with familial"						
BASE COUNT	28	a	20	c	40	g	20 t
ORIGIN							
Query Match	0.3%; Score 79.4; DB 10; Length 108;						
Best Local Similarity	84.8%; Pred. No. 0.00025;						
Matches	89;	Conservative	0;	Mismatches	16;	Indels	0; Gaps 0;
QY	20131	TCGGCTCACTGCAACCTCCGCCTCCCGCATTCAGGCCATTTCTCCTGGCTCAACTCCCGA	20190				
Db	107	TCGCCTACCACAACCTCTGCCTCTCGGTTCAAACCATTTTCTCGCTCAGCCTCCGA	48				
QY	20191	GTACTGGGACCACAGGCGCCGCCACCAACCCCAGCTAATTTTT	20235				
Db	47	GTAGCTGGGATTACAGCACCTGTCACACACGCTGGCTAATTTTT	3				
RESULT 13							
HS8IC8R/c							
LOCUS	HS8IC8R	103 bp	DNA	STS	05-SEP-1991		
DEFINITION	Human sequence tagged site 8IC8R DNA from 19q13.						
ACCESSION	X57789						
VERSION	X57789.1	GI:23938					
KEYWORDS	STS; myotonic dystrophy.						
SOURCE	human.						
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
REFERENCE	1 (bases 1 to 103) Aldridge,F.L.						
AUTHORS	Direct Submission						
TITLE	Submitted (12-FEB-1991) F.L. Aldridge, ICI Pharmaceuticals,						
JOURNAL	Alderley Park, Macclesfield, Cheshire, SK10 4TG, UK						
REFERENCE	2 (bases 1 to 103) Butler,R., Riley,J.H., Ogilvie,D.J., Anand,R., Buxton,J., Davies,J., Johnson,K. and Markham,A.F.						
AUTHORS	Two sequence-tagged sites defining the ends of a 380 kb YAC clone from 19q13						
TITLE	Nucleic Acids Res. 19 (17), 4787 (1991)						
JOURNAL MEDLINE	91367697						
COMMENT	See also X57788 for STS 8IC8L.						
FEATURES							
source	1..103 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="19q13" /germline /clone_lib="YAC library: ICI" /clone="8IC8"						
BASE COUNT	29	a	28	c	23	g	22 t 1 others
ORIGIN							




```
RESULT 2
ID X12095 standard; DNA; 108 BP.
AC X12095;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment TIGR-A003M18a.
KW polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PR 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHEED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI: 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1: Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
CC Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;
SQ
```

```
Query Match 0.2%; Score 68.6; DB 1; Length 108;
Best Local Similarity 82.8%; Pred. No. 0.015;
Matches 77; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

QY 10310 TGTATTTTGTAGACGGGTTTTCACATGTTGGCCAGGCTGCTCTCAAACTCCTGAC 10369
DB 1 TGTCTTTTGTAGAGATGAGGTTTTCCTCTGTTGGCCAGGATGCTCTCGAACTCCTGAC 60

QY 10370 CTCAGGTGATCCACCTCGCTCGGCTCCCAAAA 10402
DB 1 TTTCAAGTGTATCGCTGCTCGCTTGGCCCTCCCAAAA 93

RESULT 3
ID T24892 standard; cDNA to mRNA; 100 BP.
AC T24892;
DT 05-NOV-1996 (first entry)
DE Human gene signature HUMGS06998.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PR 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
```

```
Query Match 0.2%; Score 63.2; DB 1; Length 100;
Best Local Similarity 77.9%; Pred. No. 0.1;
Matches 74; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 28738 ATCACTTGAACCTGGAGGCGAGAGTTGTCAGTGAGGGAGATGGCCACTGCACCTCCAG 28797
DB 2 ATCGCTTGAACCTGGAGGCGAGATTTTGCATNAGCTGAGATTGCACCTTGCACCTCCG 61

QY 28798 CCTGAGCAACACAGCGAGACTCTGTCTCAAAAAA 28832
DB 62 CCTGGGTGACAGAGTGACACTCTGTGTGAACAAA 96

RESULT 4
ID T25009 standard; cDNA to mRNA; 108 BP.
AC T25009;
DT 07-NOV-1996 (first entry)
DE Human gene signature HUMGS07131.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PR 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1748; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
CC Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;
SQ
```

```
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1720; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
CC Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;
SQ
```

```
Query Match 0.2%; Score 63.2; DB 1; Length 100;
Best Local Similarity 77.9%; Pred. No. 0.1;
Matches 74; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 28738 ATCACTTGAACCTGGAGGCGAGAGTTGTCAGTGAGGGAGATGGCCACTGCACCTCCAG 28797
DB 2 ATCGCTTGAACCTGGAGGCGAGATTTTGCATNAGCTGAGATTGCACCTTGCACCTCCG 61

QY 28798 CCTGAGCAACACAGCGAGACTCTGTCTCAAAAAA 28832
DB 62 CCTGGGTGACAGAGTGACACTCTGTGTGAACAAA 96

RESULT 4
ID T25009 standard; cDNA to mRNA; 108 BP.
AC T25009;
DT 07-NOV-1996 (first entry)
DE Human gene signature HUMGS07131.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PR 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1748; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
```


PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

Claim 1: Page 1720; 2245pp; Japanese.
A single-stranded DNA (or its complementary strand or the corresp.
double-stranded DNA) which comprises one of the 7837 "GS" sequences
given in T19001-T26837 and which is able to hybridise to part of
human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
sequences were obtained from 3'-directed cDNA libraries prepared
from various human tissues; synthesis of cDNA was initiated from the
3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
untranslated sequence is unique to a particular mRNA species, almost
all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
is constructed so as to reflect accurately the relative abundance of
different mRNAs in the particular tissue from which it was derived.
The appearance frequency of a given GS in a cDNA library can be
determined (esp. using primers and probes derived from the GS
sequences) as a means of diagnosing abnormal cell function or for
recognising different cell types.
Sequence 100 BP: 28 A; 22 C;
Sequence 100 BP: 25 G; 22 T;

	Query Match	0.28;	Score 59.2;	DB 1;	Length 100;
	Best Local Similarity	73.78;	Pred. No. 0.41;	26;	Indels 0; Gaps 0;
	Matches	73;	Conservative		
Qy	20071	TTTCTTTTTCTGAGACGGAGCTTTGCTCTGTCGCCAGCTGGAAATCGAGTCGGCACAATC	20130		
Db	100	TTTGTGTTGTTTCAACACAGAGTGTCACTCTGTCACCCAGCGNGGAGTGCAANGGTGCAATC	41		
Qy	20131	TCGGCTCACTGCAACCTCCGGCTCCGGGATTCAGGCCAT	20169		
Db	40	TCAGCTNATGCAAAATCTGCCTCCCAAGGTTCAAGCGAT	2		

RESULT	8	
T25009/c		
ID	T25009 standard; cDNA to mRNA; 108 BP.	
AC	T25009;	
DT	07-NOV-1996 (first entry)	
DE	Human gene signature HUMGS07131.	
KE	Gene signature; messenger RNA; mRNA; relative abundance; frequency;	
KW	human; cloning; mapping; non-biased library; diagnosis; detection;	
KS	cell typing; abnormal cell function; ss.	
QW	Homo sapiens.	
PN	WO9514772-A1.	
PD	01-JUN-1995.	
PF	11-NOV-1994; J01916.	
PR	12-NOV-1993; JP-355504.	
PA	(MATS/) MATSUBARA K.	
PK	(OKUB/) OKUBO K.	
PI	Matsubara K, Okubo K;	
WPI	95-206931/27.	
DR	Identifying gene signatures in 3'-directed human cDNA library - e.g.	
PT	for diagnosis of abnormal cell function, by preparing cDNA that	
PT	reflects relative abundance of corresp. mRNA in specific human	
PT	tissues	

Seq	Sequence	108 BP;	34 A;	31 C;	26 G;	15 T;
	Query Match		0.28;	Score 58.6;	DB 1;	Length 108;
	Best Local Similarity		71.0%;	Pred. No. 0.51;		
	Matches	76;	Conservative	0;	Mismatches	31;
					Indels	0;
					Gaps	0;
Qy	20063	TTTTTTTTTTCTTTTTCTGACGGAGTCTTGCTCTCTCGCCAGGCTGGAATGCAGTG	20122			
Db	108	TTTGTGTTGTTGTTGTTTCAACAGGCTCTTGCTCTCACTCAGGCTGGAATNCAGTG	49			
Qy	20123	GCACAACTCGGCTCACTGCARACCTCCGCCCTCCCGGATTCAGGCAT	20169			
Db	48	CGGTGACCATGGCTCACTGCAGCCTTGCCCTCAATGGGCTCAGGCAT	2			

RESULT	9	
T21566		
ID	T21566 standard; cDNA to mRNA; 87 BP.	
AC	T21566;	
DC	03-AUG-1996 (first entry)	
DE	Human gene signature HUMGS02944.	
KE	Gene signature; messenger RNA; mRNA; relative abundance; frequency;	
KW	human; cloning; mapping; non-biased library; diagnosis; detection;	
KW	cell typing; abnormal cell function; ss.	
KW	Homo sapiens.	
PN	W09514772-A1.	
PN	01-JUN-1995.	
PF	11-NOV-1994; J01916.	
PR	12-NOV-1993; JP-355504.	
PA	(MATS/) MATSUBARA K.	
PA	(OKUB/) OKUBO K.	
PI	Matsubara K. Okubo K;	
DR	WPI: 95-205931/27.	
PT	Identifying gene signatures in 3'-directed human cDNA library - e.g.	
PT	for diagnosis of abnormal cell function, by preparing cDNA that	
PT	reflects relative abundance of corresp. mRNA in specific human	
PT	tissues	
PS	Claim 1: Page 914; 2245pp; Japanese.	
CC	A single-stranded DNA (or its complementary strand or the corresp.	
CC	double-stranded DNA) which comprises one of the 7837 "GS" sequences	
CC	given in T19001-T26837 and which is able to hybridise to part of	
CC	human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)	
CC	sequences were obtained from 3'-directed cDNA libraries prepared	
CC	from various human tissues; synthesis of cDNA was initiated from the	
CC	3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-	
CC	untranslated sequence is unique to a particular mRNA species, almost	
CC	all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library	
CC	is constructed so as to reflect accurately the relative abundance of	
CC	different mRNAs in the particular tissue from which it was derived.	
CC	The appearance frequency of a given GS in a cDNA library can be	
CC	determined (esp. using primers and probes derived from the GS	
CC	sequences) as a means of diagnosing abnormal cell function or for	
CC	recognising different cell types.	
SQ	Sequence 87 BP; 35 A; 21 C; 16 G; 13 T;	

[illegible]

RESULT 10
T21566/c
ID T21566 standard: cDNA to mRNA: 87 BP.


```
Query Match          0.2%; Score 57.4; DB 1; Length 97;
Best Local Similarity 75.3%; Pred. No. 0.78;
Matches 70; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 7593 GATCTGGCCACTCGGCTCCGAGAGTGTGGATTACAGGTGTGAGCCAGCCGAGA 7652
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GATCTGCCACCTNGCGCTCCAGAGTGTGGATTACAGGCATGAGCCATGCGCCGG 60

QY 7653 CCTGGACTTGTCTCTGTTTCATCAGTCCTTC 7685
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 NCTGTACTAAGTCTTTTCTTTTAAATTCCTC 93

RESULT 13
ID X12087/c
AC X12087 standard; DNA; 100 BP.
DE 30-MAR-1999 (first entry)
KW Human biallelic polymorphic DNA fragment EST98276a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-TUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC spherocytosis, von Willebrand's disease, polycystic kidney disease, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match          0.2%; Score 57; DB 1; Length 100;
Best Local Similarity 72.7%; Pred. No. 0.9;
Matches 72; Conservative 1; Mismatches 26; Indels 0; Gaps 0;

QY 25232 GTGGCTCACCTGTAAATCCAGCACTTTGGAGGCCAAGGTAAAGCATCATTGAGGT 25291
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 99 GTGACTCACCTATAATCTTGGCACCTTTGGAGGCTTAGGAAGGAGGATTGTTTGAAC 40

QY 25292 CAGGAGTTAGAGACCACTCTGCCCAACATAGTGAACATC 25330
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 39 CAGGAGCTCAAGACCATCTCTGGGAACATAGCAAGATC 1

RESULT 14
ID X12085/c
AC X12085 standard; DNA; 100 BP.
DE 30-MAR-1999 (first entry)
KW Human biallelic polymorphic DNA fragment EST98276c.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 218; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-TUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberculous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match          0.2%; Score 57; DB 1; Length 100;
Best Local Similarity 72.7%; Pred. No. 0.9;
Matches 72; Conservative 1; Mismatches 26; Indels 0; Gaps 0;

QY 25232 GTGGCTCACCTGTAAATCCAGCACTTTGGAGGCCAAGGTAAAGCATCATTGAGGT 25291
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 99 GTGACTCACCTATAATCTTGGCACCTTTGGAGGCTTAGGAAGGAGGATTGTTTGAAC 40

QY 25292 CAGGAGTTAGAGACCACTCTGCCCAACATAGTGAACATC 25330
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 39 CAGGAGCTCAAGACCAKCTCTGGGAACATAGCAAGATC 1

RESULT 14
ID X12085/c
AC X12085 standard; DNA; 100 BP.
DE 30-MAR-1999 (first entry)
KW Human biallelic polymorphic DNA fragment EST98276c.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 218; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-TUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberculous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;
```

```
AC X12085;
DE 30-MAR-1999 (first entry)
KW Human biallelic polymorphic DNA fragment EST98276c.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 218; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-TUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberculous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 22 A; 25 C; 22 G; 30 T;

Query Match          0.2%; Score 57; DB 1; Length 100;
Best Local Similarity 72.7%; Pred. No. 0.9;
Matches 72; Conservative 1; Mismatches 26; Indels 0; Gaps 0;

QY 25232 GTGGCTCACCTGTAAATCCAGCACTTTGGAGGCCAAGGTAAAGCATCATTGAGGT 25291
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 99 GTGACTCACCTATAATCTTGGCACCTTTGGAGGCTTAGGAAGGAGGATTGTTTGAAC 40

QY 25292 CAGGAGTTAGAGACCACTCTGCCCAACATAGTGAACATC 25330
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 39 CAGGAGCTCAAGACCATCTCTGGGAACATAGCAAGATC 1

RESULT 15
ID T20927
AC T20927 standard; cDNA to mRNA; 103 BP.
DE 24-JUL-1996 (first entry)
KW Human gene signature H0MG502180.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
```

PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1: Page 758-759; 2245pp; Japanese.

CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 103 BP; 22 A; 27 C; 21 G; 31 T;

Query Match 0.2%; Score 57.2; DB 1; Length 103;
Best Local Similarity 77.3%; Pred. No. 0.84;
Matches 68; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 10254 TCCTCTCCCAAGTAGCTGGGACTACAGTGCATACCCACCGCCCTGCTAATTTTGTGA 10313

Db 13 TCACCTCCCAAGTAGCTGGGCTACAGGTGTGCCACCATGTCCAGCTGATTNGTA 72

QY 10314 TTTTGTATAGACGGGGTTTCACCATG 10341

Db 73 TTTTGTATAGACGGGGTTTCACCATG 100

Search completed: June 16, 2000, 12:11:56
Job time: 170027 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2000, 19:11:36 ; Search time 8514.64 Seconds
(without alignments)
13805.343 Million cell updates/sec

Title: US-08-852-495C-1_COPY_112000_141000
Perfect score: 29001
Sequence: 1 TTTTCCACTCTTCTTCAG.....AGAGTGTTCACCTCTAGGA 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues
Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : EST:
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
105: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result	%	Query
SUMMARIES		

No.	Score	Match	Length	DB	ID	Description
1	93.2	0.3	106	37	AA703692	AA703692 ag81a10.r
2	91.4	0.3	109	30	AA243009	AA243009 zr25h02.s
3	90.4	0.3	106	30	AA250812	AA250812 zs06a05.s
c	88.6	0.3	103	84	B48914	B48914 RPC111-4A12
5	88.2	0.3	101	39	AA835205	AA835205 ak64h01.s
c	88.2	0.3	110	30	AA244245	AA244245 nc07a04.s
c	87.8	0.3	101	33	AA381369	AA381369 EST94442
c	87.8	0.3	108	84	B65160	B65160 CIT-HSP-201
c	86.6	0.3	102	36	AA654562	AA654562 nt75f10.s
c	86.2	0.3	103	108	AQ582186	AQ582186 RPC1-11-4
c	86.2	0.3	107	35	AA565533	AA565533 nk42b11.s
c	86.4	0.3	107	39	AA828124	AA828124 od71a07.s
c	86.6	0.3	110	94	AQ003188	AQ003188 RPC111-1D
c	86.2	0.3	102	36	AA654562	AA654562 nt75f10.s
c	85.4	0.3	103	108	AQ535244	AQ535244 RPC1-11-3
c	85.2	0.3	109	94	AQ028426	AQ028426 CIT-HSP-2
c	85.2	0.3	110	30	AA244245	AA244245 nc07a04.s
c	85.2	0.3	110	39	AA897366	AA897366 am06h02.s
c	84.4	0.3	105	21	T94466	T94466 ve35b02.r1
c	84.6	0.3	107	103	AQ240182	AQ240182 CIT-HSP-2
c	84.2	0.3	106	108	AQ544957	AQ544957 CITBI-EI-
c	84.4	0.3	110	106	AQ386882	AQ386882 RPC111-13
c	83.8	0.3	103	94	AQ028649	AQ028649 CIT-HSP-2
c	83.8	0.3	103	108	AQ535244	AQ535244 RPC1-11-3
c	83.6	0.3	106	63	AI991750	AI991750 wt48e01.x
c	83.6	0.3	106	63	AI991750	AI991750 wt48e01.x
c	83.4	0.3	109	84	B17434	B17434 345K2.IVB C
c	83.6	0.3	109	84	B17434	B17434 345K2.IVB C
c	83.2	0.3	96	92	AQ936334	AQ936334 RPC1-11-S
c	83.0	0.3	101	33	AA381369	AA381369 EST94442
c	83.2	0.3	104	105	AQ321855	AQ321855 RPC111-11
c	82.8	0.3	103	38	AA807640	AA807640 nx08b05.s
c	82.2	0.3	103	108	AQ584425	AQ584425 RPC1-11-4
c	82.2	0.3	105	109	AQ637292	AQ637292 RPC1-11-4
c	82.4	0.3	108	84	B32951	B32951 HS-1016-A1-
c	81.8	0.3	106	30	AA250812	AA250812 zs06a05.s
c	82.0	0.3	106	38	AA812141	AA812141 ob48h02.s
c	82.0	0.3	106	50	AI700000	AI700000 tt36a10.x
c	81.6	0.3	104	105	AQ321855	AQ321855 RPC111-11
c	81.6	0.3	105	30	AA218889	AA218889 zq15d04.s
c	81.0	0.3	105	28	AA078003	AA078003 7H12D08 C
c	81.2	0.3	110	106	AQ386882	AQ386882 RPC111-13
c	80.6	0.3	107	33	AA385808	AA385808 EST99495
c	80.2	0.3	101	94	AQ076649	AQ076649 CIT-HSP-2
c	80.2	0.3	101	105	AQ260734	AQ260734 CITBI-EI-

ALIGNMENTS

RESULT 1
AA703692
LOCUS ag81a10.r1 StrataGene hnt neuron (#937233) Homo sapiens cDNA clone IMAGE:1140858 5' similar to contains Alu repetitive element; mRNA sequence.
DEFINITION
ACCESSION AA703692
VERSION AA703692.1 GI:2713610
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 106)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1397630.

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 53.
Location/Qualifiers
FEATURES
source 1..106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1140858"
/clone_lib="StrataGene hnt neuron (#937233)"
/dev_stage="hnt neurons"
/lab_host="SOLR (kanamycin resistant)"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt.
Differentially, post mitotic hnt neurons. Average insert size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATCGGCACGAG 3' -3' adaptor sequence: 3' CTCAGCTTTTCTTTTCTTTT 3' 29 t
BASE COUNT 19 a 29 c 29 g 29 t
ORIGIN
Query Match 0.3%; Score 93.2; DB 37; Length 106;
Best Local Similarity 92.5%; Pred. No. 0.2;
Matches 98; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 20240 TTTTGTAGACAGAGGTTTCACCGTGTAGCCGGGATGCTCGATCTCCTGACCTCA 20299
|||||
Db 1 TTTTGTAGACAGAGGTTTCACCGTGTAGCCGGGATGCTCGATCTCCTGACCTCG 60
QY 20300 TGATCTGCCCCCTCAGCCTCCCAAGTGTAGGATCAGAGCGCATG 20345
|||||
Db 61 TGATCTGCCCCCTCAGCCTCCCAAGTGTAGGATCAGAGCGCATG 106
RESULT 2
AA243009 109 bp mRNA EST 11-MAR-1998
LOCUS zr25h02.s1 StrataGene NT2 neuronal precursor 937230 Homo sapiens
DEFINITION cDNA clone IMAGE:664467 3' similar to contains Alu repetitive element; contains element LTR1 repetitive element ;, mRNA sequence.
ACCESSION AA243009
VERSION AA243009.1 GI:1873869
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 109)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Dec 3, 1996 this sequence version replaced gi:1126869.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 1127 Std Error: 0.00
Seq primer: -4m13 fwd. ET from Amersham
High quality sequence stop: 102.

FEATURES
source

```
Location/Qualifiers
1. .109
/organism="Homo sapiens"
/db_xref="GDB:5426481"
/db_xref="taxon:9606"
/clone="IMAGE:664467"
/tissue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: brain; Vector: phuescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Uninduced, exponentially growing neuroepithelial
cells (Ntera-2/cl.D1). Average insert size: 1.0 kb;
Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCAGGAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3"
BASE COUNT      19 a      30 c      30 g      30 t
ORIGIN
Query Match      0.3%; Score 91.4; DB 30; Length 109;
Best Local Similarity 89.9%; Pred. No. 0.31; Indels 0; Gaps 0;
Matches 98; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 20237 GYATTTTGTAGACAGCGGTTTACCCTGTAGCCGGGATGGTCTCGATCTCCTGACC 20296
|||||
Db 1 GYATTTTGTAGACAGCGGTTTACCCTGTAGCCGGGATGGTCTCGATCTCCTGACC 60
|||||

QY 20297 TCATGATCGCCACCTGACGCTCCCAAGTGTAGGATCAGCAGGATG 20345
|||||
Db 61 TCGTGATCGCCACCTGACGCTCCCAAGTGTAGGATCAGCAGGATG 109
|||||

RESULT 3
AA250812      106 bp      mRNA      EST      15-AUG-1997
LOCUS      zs06a05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:684368 3'
DEFINITION      similar to contains Alu repetitive element; contains element MER22
repetitive element ; mRNA sequence.
ACCESSION      AA250812
VERSION
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 106)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      On Sep 12, 1996 this sequence version replaced gi:1407356.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 537 Std Error: 0.00
Seq primer: -41m13 fwd. Et from Amersham
High quality sequence stop: 87.
Location/Qualifiers
1. .106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:684368"
/tissue_type="NCI_CGAP_GCB1"
/lab_host="DH108"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
```

FEATURES
source

```
BASE COUNT      20 a      28 c      31 g      27 t
ORIGIN
Query Match      0.3%; Score 90.4; DB 30; Length 106;
Best Local Similarity 94.0%; Pred. No. 0.4;
Matches 94; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 10321 TAGAGACGGGGTTTACCATGTTGGCCAGGCTGCTCAAACTCTTGGAGCTCAGGTGATC 10380
|||||
Db 2 TAGAGACGGGGTTTACCATGTTGGCCAGGCTGCTCAAACTCTTGGAGCTCAGGTGATC 61
|||||

QY 10381 CACCTCGCTCGCCCTCCCAAAATGCTGAGATTACAGGTGT 10420
|||||
Db 62 CACTTGCCTTGGCCCTCCCAAAATGCTGAGATTACAGGTGT 101
|||||

RESULT 4
B48914/c
LOCUS      B48914      103 bp      DNA      GSS      08-APR-1999
DEFINITION      RPC111-4A12.TP RPCI-11 Homo sapiens genomic clone RPCI-11-4A12,
genomic survey sequence.
ACCESSION      B48914
VERSION      B48914.1 GI:2601151
KEYWORDS      GSS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 103)
AUTHORS      Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venner,J.C.
TITLE      Use of BAC End Sequences for Sequence-Ready Map Building
JOURNAL      Unpublished (1997)
COMMENT      Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1. .103
/organism="Homo sapiens"
/db_xref="GDB:7501163"
/db_xref="taxon:9606"
/clone="RPCI-11-4A12"
/tissue_type="RPCI-11"
/lab_host="Male"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC Library"
BASE COUNT      30 a      28 c      30 g      15 t
ORIGIN
```



```

RESULT 7
LOCUS AA381369/c
DEFINITION EST94442 Activated T-cells 1 Homo sapiens cDNA 5' end similar to
ACCESSION AA381369
VERSION AA381369.1 GI:2033689
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS 1 (bases 1 to 101)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Whai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Melssner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
TITLE Initial assessment of human gene diversity and expression patterns
JOURNAL based upon 83 million nucleotides of cDNA sequence
MEDLINE Nature 377 (6547 Suppl), 3-174 (1995)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1407448.
Other_ESTs: THC170052
Contact: kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tldb/hgi/hgi.html)
Seq primer: M13 Reverse.
FEATURES
Location/Qualifiers
1..101
/organism="Homo sapiens"
/db_xref="ATCC (Inhost):185728"
/db_xref="taxon:9606"
/clone_lib="Activated T-cells 1"
/cell_type="T-lymphocyte"
/dev_stage="adult"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 18 a 36 c 20 g 25 t 2 others
ORIGIN
Query Match 0.3%; Score 87.8; DB 33; Length 101;
Best Local Similarity 91.1%; Pred. No. 0.78;
Matches 92; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 13186 AAAATTAGCTGGGTGGTGGCGGCACCTGTAATCCAGCTAATCAGGAGCTGAGGCA 13245
Db 101 AAAATTAGCTGGGAGTGGTGGCGGCGCCCTGTATCCAGCTACTCAGGAGGCTGAGGCA 42
QY 13246 GGAGATCGCTTGAAACCCAGGAGGAGAGGTTGCAGTGAGC 13286
Db 41 GGANAATTGCTTCAACCCAGGAGGCGGAGGTTGCAATGAGC 1

```

```

RESULT 8
LOCUS B65160
DEFINITION CIT-HSP-2017G2.TRB CIT-HSP Homo sapiens genomic clone 2017G2,
ACCESSION B65160
VERSION B65160.1 GI:2639138
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS 1 (bases 1 to 108)
Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
JOURNAL Building
COMMENT Unpublished (1997)
Other_GSSs: CIT-HSP-2017G2.TFB
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES
Location/Qualifiers
1..108
/organism="Homo sapiens"
/db_xref="GDB:7043860"
/db_xref="taxon:9606"
/clone="2017G2"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT 26 a 27 c 34 g 21 t
ORIGIN
Query Match 0.3%; Score 87.8; DB 84; Length 108;
Best Local Similarity 88.8%; Pred. No. 0.76;
Matches 95; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 25365 TGGTGGCATGTGCTGTATCCAGCTACTCAGGAGCTGAGGAGAAATGCTTGAA 25424
Db 1 TGGTGGCATGCGCCTGTATCCAGCTACTCAGGAGCTGAGGAGAAATGCTTGAA 60
QY 25425 CCAGAGGCGGAGGTGTCAGTGGCCGAAATCGCGCACTGCACTC 25471
Db 61 CCGGGAGGTGGAGTTGTCAGTGAGCCAGATCATACCTGCACAC 107

```

```

RESULT 9
LOCUS AA654562/c
DEFINITION nt75f10.s1 NCI CGAP.Pr3 Homo sapiens cDNA clone IMAGE:1204363
ACCESSION AA654562
VERSION AA654562.1 GI:2590716
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS 1 (bases 1 to 102 bp mRNA
nt75f10.s1 NCI CGAP.Pr3 Homo sapiens cDNA clone IMAGE:1204363
similar to contains Alu repetitive element; contains element MER22
repetitive element ;, mRNA sequence.
ACCESSION AA654562
VERSION AA654562.1 GI:2590716
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates: Catarrhini; Hominidae; Homo.
 1 (bases 1 to 102)
 NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 On Sep 12, 1996 this sequence version replaced gi:1393451.
 Contact: Robert Strausberg, Ph.D.
 Tel.: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui,
 M.D., Michael Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: David B. Krizman, Ph.D.
 cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40m13 fwd. ET from Amersham.

```

FEATURES
source
1. .102
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1204363"
/clone_lib="NCI_CGAP_Pr3"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/notes="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected cells
histologically-determined to be fully malignant prostate
cancer cells. Double-stranded cDNA was ligated to EcoRI
adaptors, 5 cycles of PCR applied to the cDNA with an
adaptor-specific primer, and the resulting PCR product
subcloned into pAMP10 by the UDG-cloning method (Life
Technologies). Average insert size is 600 bp. NOTE: Not
directionally cloned. This library was constructed by
David Krizman."

```

BASE COUNT	22 a	32 c	27 g	21 t
ORIGIN				
Query Match		0.3%	Score 86.6;	DB 36; Length 102;
Best Local Similarity		91.1%;	Pred. NO. 1.1;	
Matches 92; Conservative		0; Mismatches	9; Indels	0; Gaps 0;

QY	13071	GCTCATCTCTGTAAATACAGACAC	TTGGGAGGCCGATGTGGGTGGATCACC	CTGAGGTCAG	13130
Db	101	GCTCACCTCTGTAAATCCAGACAC	TTTGGGAGGCCGATGTGGGTGGATCACC	CTGAGGTCAG	42
QY	13131	GAGTTTGAGACCGACAGCTGGCC	AACATGGTGAACCTCATCT		13171
Db	41	GAGTTTGAGACCGACAGCTGGCC	AACATGGTGAACCTCATCT		1

RESULT	10
AQ582186	
LOCUS	AQ582186 103 bp DNA GSS 07-JUN-1999
DEFINITION	RPCI-11-451A15.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-451A15, genomic survey sequence.
ACCESSION	AQ582186
VERSION	AQ582186.1 GI:5009296
KEYWORDS	GSS.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 103)
REFERENCE	Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P., and

Venter, J. C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
Unpublished (1997)
On Feb 19, 1999 this sequence version replaced gi:4146076.
Other GSSs: RPCI-11-451A15.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.
See primer: sp6

FEATURES	source	Class. cont ends.	Location/Qualifiers
			1. .103
			/organism="Homo sapiens"
			/db_xref="GDB:7672814"
			/db_xref="taxon:9606"
			/clone="RPC1-11-451A15"
			/clone_lib="RPC1-11"
			/sex="Male"
			/cell_type="Lymphocytes"
			/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; RPOC11 Human Male BAC Library"
BASE COUNT	19 a	36 c	25 g
ORIGIN			22 t
			1 others

BASE COUNT	19 a	36 c	25 g	22 t	1 others
ORIGIN					
Query Match			0.3%	Score 86;	DB 108;
Best Local Similarity			89.3%;	Pred. No. 1.2;	Length 103;
Matches 92:	Conservative	0;	Mismatches	11;	Indels 0;
					Gaps 0;

	Qy	20112	GGAAATGAGTGGGCACAAATCTCGGCTCACTGCAACCTCGCGCTCCGGGATTCAAGCCATTTC	20171
	Db	1	GGATGGNGTGGCACAATCTCGGCTCACTGCAACCTCTGCTCCAGATTCAAGCGATTT	60
	Qy	20172	TCCTGCTCTAACCTCCGAGTAGCTGGGACCACACAGGCGCGCGC	20214
	Db	61	TCCTGCTCTACGCTCCGAGTAGCTGGAGCTACAGACGCTTGC	103

[illegible]


```
BASE COUNT      22 a 27 c 26 g 35 t
ORIGIN

/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC Library"
Query Match      0.3%; Score 86.6; DB 94; Length 110;
Best Local Similarity 87.2%; Pred. No. 1;
Matches 95; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 20232 TTTTGTATTTTAGTAGACAGAGGTTTCACCGCTGTAGCGGGATGTCGTCGATCC 20291
|||||
Db 2 TTTTGTATTTTAGTAGACAGGTTTACCATGTTGGCCAGGATGTCGCCGATCTCT 61
|||||

QY 20292 TGACCTCATGATCTGCCACCTCAGCCCTCCCAAGTGTAGATCAGACAG 20340
|||||
Db 62 TGACCTCATGATCCACCTGCCCGCCAGCCCTCCCAAGTGTGGGATTACAG 110
|||||

RESULT 14
AA654562
LOCUS      AA654562      102 bp      mRNA      EST      04-NOV-1997
DEFINITION nt75f10.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1204363
similar to contains Alu repetitive element;contains element MER22
repetitive element ;, mRNA sequence.
ACCESSION  AA654562
VERSION     AA654562.1 GI:2590716
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 102)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1393451.
Contact: Robert Strausberg, Ph.D.
Tel.: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui,
M.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.W.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40ml3 fwd. ET from Amersham.
FEATURES             Location/Qualifiers
     source           1..102
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /clone="IMAGE:1204363"
     /clone.lib="NCI_CGAP_Pr3"
     /sex="Male"
     /dev_stage="45 years old"
     /lab_host="DH10B"
     /note="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected cells
histologically-determined to be fully malignant prostate
cancer cells. Double-stranded cDNA was ligated to EcoRI
adaptors, 5 cycles of PCR applied to the cDNA with an
adaptor-specific primer, and the resulting PCR product
subcloned into pAMP10 by the UDG-cloning method (Life
Technologies). Average insert size is 600 bp. NOTE: Not
directionally cloned. This library was constructed by
David Krizman."
BASE COUNT      22 a 32 c 27 g 21 t
```

```
ORIGIN

Query Match      0.3%; Score 86; DB 36; Length 102;
Best Local Similarity 90.2%; Pred. No. 1.2;
Matches 92; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 10324 AGACGGGTTTCACCATGTTGGCCAGGCTGGTCTCAAACTCTCTGACCTCAGGTGATCCAC 10383
|||||
Db 1 AGACGAGGTTTCACCATGTTGGCCAGTCTGTCTCAAACTCCGGACCTCAGGTAATCCGC 60
|||||

QY 10384 CTGCCCTCGCCCTCCCAAAATGCTGAGATTACAGGTGTGAGGCC 10425
|||||
Db 61 CCACCTCGCCCTCCCAAGTGTCTGGGATTACAGGAGTGAGCC 102
|||||

RESULT 15
AQ535244/c
LOCUS      AQ535244      103 bp      DNA      GSS      18-MAY-1999
DEFINITION RPCI-11-317H22.TV RPCI-11 Homo sapiens genomic clone
RPCI-11-317H22, genomic survey sequence.
ACCESSION  AQ535244
VERSION     AQ535244.1 GI:4846934
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 103)
AUTHORS     Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and
Venter, J.C.
TITLE       Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
JOURNAL     Unpublished (1997)
COMMENT     Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: 17
Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..103
     /organism="Homo sapiens"
     /db_xref="GDB:7621533"
     /db_xref="taxon:9606"
     /clone="RPCI-11-317H22"
     /clone.lib="RPCI-11"
     /sex="Male"
     /cell_type="Lymphocytes"
     /note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC Library"
BASE COUNT      31 a 27 c 27 g 18 t
ORIGIN

Query Match      0.3%; Score 85.4; DB 108; Length 103;
Best Local Similarity 89.3%; Pred. No. 1.4;
Matches 92; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 7522 TTTTGTATTTTAGTAGACAGAGTGGGTTTCACCATGTTGGCCAGACTGCTCTCAACTCC 7581
|||||
Db 103 TTTTGTATTTATACACAGACGAGGGGTTTCACCATGTTGGCCAGGCTGCTCTCAACTCC 44
|||||

QY 7582 TGGCCTCAAGTGATCTGGCCACCTCGGCTCCCGAAGTGCTGG 7624
```

Db 43 TGACCTCAAGTCATCTGCCCCGCTTTGGCCTCCCAAAGTCTGG 1

Search completed: June 16, 2000, 03:46:29
Job time: 140956 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2000, 00:17:42 ; Search time 372.34 Seconds
(without alignments)
10124.339 Million cell updates/sec

Title: US-08-852-495C-1_COPY_112000_141000
Perfect score: 29001
Sequence: 1 TTTTCCACCTCTTCTTCAG.....AGACTGTTGACCTCTAGGA 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78.4	0.3	105	4	US-08-481-658B-65
2	78.4	0.3	105	4	US-08-477-504A-65
3	78.4	0.3	105	4	US-08-486-756A-65
4	78.4	0.3	105	4	US-08-485-862B-65
5	78.4	0.3	105	5	US-08-787-739-65
6	68	0.2	105	4	US-08-481-658B-65
7	68	0.2	105	4	US-08-477-504A-65
8	68	0.2	105	4	US-08-486-756A-65
9	68	0.2	105	4	US-08-485-862B-65
10	68	0.2	105	5	US-08-787-739-65
11	60.4	0.2	78	3	US-08-454-557C-70
12	60.4	0.2	78	4	US-08-340-426D-70
13	60.4	0.2	78	4	US-08-450-673C-70
14	60.4	0.2	78	6	PCT-US95-17111A-70
15	58.8	0.2	78	3	US-08-454-557C-70
16	58.8	0.2	78	4	US-08-340-426D-70
17	58.8	0.2	78	4	US-08-450-673C-70
18	58.8	0.2	78	6	PCT-US95-17111A-70
19	59	0.2	84	3	US-08-454-557C-91
20	59	0.2	84	4	US-08-340-426D-91
21	59	0.2	84	4	US-08-450-673C-91
22	59	0.2	84	6	PCT-US95-17111A-91
23	53.2	0.2	76	3	US-08-454-557C-69
24	53.2	0.2	76	4	US-08-340-426D-69
25	53.2	0.2	76	4	US-08-450-673C-69
26	53.2	0.2	76	6	PCT-US95-17111A-69
27	53.4	0.2	85	3	US-08-454-557C-92

28	53.4	0.2	85	4	US-08-340-426D-92	Sequence 92, Appl
29	53.4	0.2	85	4	US-08-450-673C-92	Sequence 92, Appl
30	53.4	0.2	85	6	PCT-US95-17111A-92	Sequence 92, Appl
31	51.8	0.2	84	3	US-08-454-557C-91	Sequence 91, Appl
32	51.8	0.2	84	4	US-08-340-426D-91	Sequence 91, Appl
33	51.8	0.2	84	4	US-08-450-673C-91	Sequence 91, Appl
34	51.8	0.2	84	6	PCT-US95-17111A-91	Sequence 91, Appl
35	50.4	0.2	60	3	US-08-454-557C-57	Sequence 57, Appl
36	50.4	0.2	60	4	US-08-340-426D-57	Sequence 57, Appl
37	50.4	0.2	60	4	US-08-450-673C-57	Sequence 57, Appl
38	50.4	0.2	60	6	PCT-US95-17111A-57	Sequence 57, Appl
39	50	0.2	76	3	US-08-454-557C-69	Sequence 69, Appl
40	50	0.2	76	4	US-08-340-426D-69	Sequence 69, Appl
41	50	0.2	76	4	US-08-450-673C-69	Sequence 69, Appl
42	50	0.2	76	6	PCT-US95-17111A-69	Sequence 69, Appl
43	50	0.2	83	4	US-08-481-658B-66	Sequence 66, Appl
44	50	0.2	83	4	US-08-477-504A-66	Sequence 66, Appl
45	50	0.2	83	5	US-08-787-739-66	Sequence 66, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-65
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauger, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.3%; Score 78.4; DB 4; Length 105;
Best Local Similarity 84.6%; Pred. No. 2.8e-08;

Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 19437 TTTTGTATTTTCTAGTAAGAGGGGTTTCACCATGTTGGTCAGGCTGTCTCCAACTCC 19496

Db 2 TTTTACATCTTTAGTAGAGACAGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy 19497 TGACCTCATGATCTGCCACCTTGGCCCTCCCAAAAGTGTGGGAT 19540

Db 62 TGACCTTGTGATCCACGAGCTCGGCCCTCCCAAAAGTGTGGGAT 105

RESULT 2

US-08-477-504A-65

; Sequence 65, Application US/08477504A

; Patent No. 5972353

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,504A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3D

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-477-504A-65

Query Match 0.3%; Score 78.4; DB 4; Length 105;

Best Local Similarity 84.6%; Pred. No. 2.8e-08;

Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 19437 TTTTGTATTTTCTAGTAAGAGGGGTTTCACCATGTTGGTCAGGCTGTCTCCAACTCC 19496

Db 2 TTTTACATCTTTAGTAGAGACAGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy 19497 TGACCTCATGATCTGCCACCTTGGCCCTCCCAAAAGTGTGGGAT 19540

Db 62 TGACCTTGTGATCCACGAGCTCGGCCCTCCCAAAAGTGTGGGAT 105

RESULT 3

US-08-486-756A-65

; Sequence 65, Application US/08486756A

; Patent No. 5981711

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,756A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-486-756A-65

Query Match 0.3%; Score 78.4; DB 4; Length 105;

Best Local Similarity 84.6%; Pred. No. 2.8e-08;

Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 19437 TTTTGTATTTTCTAGTAAGAGGGGTTTCACCATGTTGGTCAGGCTGTCTCCAACTCC 19496

Db 2 TTTTACATCTTTAGTAGAGACAGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy 19497 TGACCTCATGATCTGCCACCTTGGCCCTCCCAAAAGTGTGGGAT 19540

Db 62 TGACCTTGTGATCCACGAGCTCGGCCCTCCCAAAAGTGTGGGAT 105

RESULT 4

US-08-485-862B-65

; Sequence 65, Application US/08485862B

; Patent No. 5989838

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

```

; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,862B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-485-862B-65

Query Match 0.3%; Score 78.4; DB 4; Length 105;
Best Local Similarity 84.6%; Pred. No. 2.8e-08;
Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 19437 TTTTGTATTTTAGTAAGACGGGGTTTCACAGTGTGGTCAGGCTGCTCCAACTCC 19496
Db 2 TTTTACATCTTTAGTAGACAGGAGGTTTACCATATGGCCAGGCTGCTCTCAAACTCC 61

QY 19497 TGACCTCATGATCGCCACCTGGCCCTCCCAAAGTGTGGGAT 19540
Db 62 TGACCTTGTGATCCACGAGCTGGCCCTCCCAAAGTGTGGGAT 105

RESULT 5
US-08-787-739-65
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481.658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260.190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.2%; Score 68; DB 4; Length 105;
Best Local Similarity 84.0%; Pred. No. 5e-06;
Matches 89; Conservative 0; Mismatches 15; Indels 2; Gaps 1;
QY 13084 ATACCAGCAGCTTTGGAGGCCGATGTGGTGATCAGCTGAGGTGAGGAGTTTGAGACCA 13143
DB 105 ATCCAGCAGCTTTGGAGGCCGATGTGGTGATCAGCTGAGGTGAGGAGTTTGAGACCA 48
QY 13144 GACTGCCCAACATGTGGAACCTCATCTCTAGTAAATAACAAAA 13189
DB 47 GCCTGCCCAATATGTGGAACCCCTGCTCTACTAAAGATGTAAAAA 2

RESULT 7
US-08-477-504A-65/c
Sequence 65, Application US/08477504A
Patent No. 5972353
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477.504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260.190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-477-504A-65
Query Match 0.2%; Score 68; DB 4; Length 105;
Best Local Similarity 84.0%; Pred. No. 5e-06;
Matches 89; Conservative 0; Mismatches 15; Indels 2; Gaps 1;
QY 13084 ATACCAGCAGCTTTGGAGGCCGATGTGGTGATCAGCTGAGGTGAGGAGTTTGAGACCA 13143
DB 105 ATCCAGCAGCTTTGGAGGCCGATGTGGTGATCAGCTGAGGTGAGGAGTTTGAGACCA 48
QY 13144 GACTGCCCAACATGTGGAACCTCATCTCTAGTAAATAACAAAA 13189
DB 47 GCCTGCCCAATATGTGGAACCCCTGCTCTACTAAAGATGTAAAAA 2
RESULT 8
US-08-486-756A-65/c
Sequence 65, Application US/08486756A
Patent No. 5981711
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486.756A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260.190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-486-756A-65

Query Match 0.2%; Score 68; DB 4; Length 105;
Best Local Similarity 84.0%; Pred. No. 5e-06;
Matches 89; Conservative 0; Mismatches 15; Indels 2; Gaps 1;

QY 13084 ATACGAGCACTTTGGAGGCCGATGTGGTGGATCACCTGAGTCAGGAGTTTGAGACCA 13143
|| |||||
DB 105 ATCCGAGCACTTTGGAGGCCGAGGCTGGTGGATCAC--AAGGTCAGGAGTTTGAGAGCA 48
|| |||||

QY 13144 GACTGCCCAACATGGTGAACCTCATCTCTAGTAAAAATACAAAA 13189
|| |||||
DB 47 GCCTGCCCAATATGGTGAACCTCTCTACTAAAGATGTAATAA 2
|| |||||

RESULT 9
US-08-485-862B-65/c
; Sequence 65, Application US/08485862B
; Patent No. 5989838
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION NUMBER: US/08/485,862B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-485-862B-65

Query Match 0.2%; Score 68; DB 4; Length 105;
Best Local Similarity 84.0%; Pred. No. 5e-06;
Matches 89; Conservative 0; Mismatches 15; Indels 2; Gaps 1;

QY 13084 ATACGAGCACTTTGGAGGCCGATGTGGTGGATCACCTGAGTCAGGAGTTTGAGACCA 13143
|| |||||
DB 105 ATCCGAGCACTTTGGAGGCCGAGGCTGGTGGATCAC--AAGGTCAGGAGTTTGAGAGCA 48
|| |||||
QY 13144 GACTGCCCAACATGGTGAACCTCATCTCTAGTAAAAATACAAAA 13189
|| |||||
DB 47 GCCTGCCCAATATGGTGAACCTCTCTACTAAAGATGTAATAA 2
|| |||||

RESULT 10
US-08-787-739-65/c
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-787-739-65

Query Match	Best Local Similarity	Score	DB 4;	DB 6;	Length	78;
Matches 67;	Conservative 0;	Mismatches 11;	Indels 0;	Gaps 0;		
Qy 19422	ACCAAGCCCGGTAATTTTGTATTTTGTAAAGACGGGGTTTACCATGTTGGTCAGG	19481				
Db 1	ACAACGCCAGCTAATTTGTATTTTGTAGATAGAGATGGGTTTCTCCATGTTTCATCAGG	60				
Qy 19482	CTGGTCTCCAACTCCTGA	19499				
Db 61	CTGGTCTCGAACTCCTGA	78				
RESULT 15						
US-08-454-557C-70/c						
Sequence 70, Application US/08454557C						
Patent No. 5830670						
GENERAL INFORMATION:						
APPLICANT: de la Monte, Suzanne						
APPLICANT: Wands, Jack R.						
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection						
TITLE OF INVENTION: of Alzheimer's Disease						
NUMBER OF SEQUENCES: 121						
CORRESPONDENCE ADDRESS:						
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.						
STREET: 1100 New York Avenue, Suite 600						
CITY: Washington						
STATE: D.C.						
COUNTRY: U.S.A.						
ZIP: 20005-3934						
COMPUTER READABLE FORM:						
MEDIUM TYPE: Floppy disk						
COMPUTER: IBM PC compatible						
OPERATING SYSTEM: PC-DOS/MS-DOS						
SOFTWARE: PatentIn Release #1.0, Version #1.25						
CURRENT APPLICATION DATA:						
APPLICATION NUMBER: US/08454,557C						
FILING DATE: 30-MAY-1995						
CLASSIFICATION: 514						
ATTORNEY/AGENT INFORMATION:						
NAME: Ludwig, Steven R.						
REGISTRATION NUMBER: 36,203						
REFERENCE/DOCKET NUMBER: 0609.3840003						
TELECOMMUNICATION INFORMATION:						
TELEPHONE: (202) 371-2600						
TELEFAX: (202) 371-2540						
INFORMATION FOR SEQ ID NO: 70:						
SEQUENCE CHARACTERISTICS:						
LENGTH: 78 base pairs						
TYPE: nucleic acid						
STRANDEDNESS: both						
TOPOLOGY: both						
US-08-450-673C-70						
Query Match	0.2%;	Score 60.4;	DB 4;	Length 78;		
Best Local Similarity	85.9%;	Pred. No. 0.00019;				
Matches 67;	Conservative 0;	Mismatches 11;	Indels 0;	Gaps 0;		
Qy 19422	ACCAAGCCCGGTAATTTTGTATTTTGTAAAGACGGGGTTTACCATGTTGGTCAGG	19481				
Db 1	ACAACGCCAGCTAATTTGTATTTTGTAGATAGAGATGGGTTTCTCCATGTTTCATCAGG	60				
Qy 19482	CTGGTCTCCAACTCCTGA	19499				
Db 61	CTGGTCTCGAACTCCTGA	78				
RESULT 14						
PCT-US95-17111A-70						
Sequence 70, Application PC/TUS9517111A						
GENERAL INFORMATION:						
APPLICANT: de la Monte, Suzanne						
APPLICANT: Wands, Jack R.						
TITLE OF INVENTION: Neural Thread Protein Gene Expression and						
TITLE OF INVENTION: Detection of Alzheimer's Disease						
NUMBER OF SEQUENCES: 121						
CORRESPONDENCE ADDRESS:						
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.						
STREET: 1100 New York Avenue, Suite 600						
CITY: Washington						
STATE: D.C.						
COUNTRY: U.S.A.						
ZIP: 20005-3934						
COMPUTER READABLE FORM:						
MEDIUM TYPE: Floppy disk						
COMPUTER: IBM PC compatible						
OPERATING SYSTEM: PC-DOS/MS-DOS						

Search completed: June 16, 2000, 08:52:43
Job time: 158293 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2000, 08:46:10 ; Search time 17973.3 Seconds
(without alignments)
-1569.657 Million cell updates/sec

Title: US-08-852-495C-l_COPY_140000_169000
Perfect score: 29001
Sequence: 1 CCCTCCAATCCATATGCAC.....TGTATACACAGAAAGTTACC 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_bal.*
- 2: gb_ba2.*
- 3: gb_om.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_pl1.*
- 8: gb_pl2.*
- 9: gb_pr1.*
- 10: gb_pr2.*
- 11: gb_pr3.*
- 12: gb_ro.*
- 13: gb_sts.*
- 14: gb_sy.*
- 15: gb_un.*
- 16: gb_vl.*
- 17: em_fun.*
- 18: em_hum1.*
- 19: em_hum2.*
- 20: em_in.*
- 21: em_om.*
- 22: em_or.*
- 23: em_ov.*
- 24: em_pat.*
- 25: em_ph.*
- 26: em_pl.*
- 27: em_ro.*
- 28: em_sts.*
- 29: em_sy.*
- 30: em_un.*
- 31: em_vl.*
- 32: gb_htg1.*
- 33: gb_htg2.*
- 34: gb_in1.*
- 35: gb_in2.*
- 36: em_ba1.*
- 37: em_ba2.*
- 38: em_hum3.*
- 39: em_hum4.*
- 40: gb_pr4.*
- 41: gb_htg3.*
- 42: gb_htg4.*
- 43: gb_htg5.*
- 44: gb_htg6.*

- 45: gb_htg7.*
- 46: em_htg1.*
- 47: em_htg2.*
- 48: em_htg3.*
- 49: em_hum5.*
- 50: gb_pl3.*
- 51: gb_pr5.*
- 52: gb_htg8.*
- 53: gb_htg9.*
- 54: gb_htg10.*
- 55: gb_htg11.*
- 56: gb_htg12.*
- 57: gb_htg13.*
- 58: gb_htg14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	85.8	0.3	108	11	HSU67803 U67803 Human small
2	85.2	0.3	108	10	HSLDLRN2 X05250 Human LDL-r
3	80.6	0.3	107	9	HUMALCE162 M87924 Human carc1
4	79.2	0.3	104	9	HUMALCE272 M87899 Human carc1
5	77.6	0.3	108	10	HSLDLRN2 X05250 Human LDL-r
6	77.8	0.3	108	11	HSU67804 U67804 Human small
7	76	0.3	103	9	HUMALCE221 M87896 Human carc1
8	76.2	0.3	108	11	HSU67808 U67808 Human small
9	75.4	0.3	103	13	HS8IC8R X57789 Human sequ
10	75	0.3	97	9	HUMLDLRA1 M14178 Human low d
11	75	0.3	107	9	HUMALCE162 M87924 Human carc1
12	74.4	0.3	103	13	HS8IC8R X57789 Human sequ
13	74.6	0.3	108	10	HSLDLRD1 X05249 Human LDL-r
14	74.6	0.3	108	10	HSLDLRD2 X05251 Human LDL-r
15	74.4	0.3	110	11	HSU67807 U67807 Human small
16	74	0.3	91	13	HUMUT8164A L30244 Human STS U
17	72.4	0.2	101	10	S79560 M87896 Human carc1
18	72.4	0.2	103	9	HUMALCE221 X05248 Human LDL-r
19	72.4	0.2	108	10	HSLDLI12 G43535 WIAF-2393-S
20	72.6	0.3	108	13	G43535 M87900 Human carc1
21	72.4	0.2	110	9	HUMALCE43 M14180 Human low d
22	72	0.2	97	9	HUMLDLRA2 U67807 Human small
23	71.8	0.2	110	11	HSU67807 L31299 Human STS U
24	71.4	0.2	97	9	HUMLDLRDJ M14179 Human famil
25	71	0.2	97	9	HUMLDLRA2 M14180 Human low d
26	70.8	0.2	100	13	HUMUT931A G32655 A009L30 Hum
27	71	0.2	105	13	G32655 M14178 Human low d
28	70.2	0.2	97	9	HUMLDLRA1 U67806 Human small
29	69.8	0.2	107	11	HSU67806 D16965 Human HepG2
30	69.4	0.2	108	9	HUMDID03M5 L30176 Human STS U
31	68.4	0.2	95	13	HUMUT8002B L30176 Human STS U
32	68.4	0.2	95	13	HUMUT8002B M14179 Human famil
33	68.6	-0.2	97	9	HUMLDLRDJ X05249 Human LDL-r
34	68.6	0.2	108	10	HSLDLRD1 X05251 Human LDL-r
35	68.6	0.2	108	10	HSLDLRD2 M15365 Human low d
36	67.6	0.2	90	9	HUMLDLRM G43535 WIAF-2393-S
37	67.8	0.2	108	13	G43535 D45223 Human GALNS
38	67.2	0.2	100	9	HUNGALNSA X66361 H.sapiens m
39	66.8	0.2	95	10	HSSTHPKIB X91545 H.sapiens D
40	66	0.2	100	10	HSLAS27 L31299 Human STS U
41	66	0.2	100	13	HUMUT931A K03555 Human low d
42	65.6	0.2	90	9	HUMLDLRFL S79560 HRX (intron
43	65.8	0.2	101	10	S79560 AR051521 Sequence
44	65.4	0.2	84	5	AR051521 G32906 A009W09 Hum
45	65.6	0.2	102	13	G32906

ALIGNMENTS

```
RESULT 1
HSU67803      108 bp      RNA      PRI      01-AUG-1997
LOCUS        Human small cytoplasmic Alu transcript.
DEFINITION
ACCESSION    U67803
VERSION      U67803.1 GI:2289917
KEYWORDS     Alu.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
              Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 108)
AUTHORS     Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE       cDNAs derived from primary and small cytoplasmic Alu (scAlu)
            transcripts
JOURNAL      J. Mol. Biol. 271 (2), 222-234 (1997)
MEDLINE      97415756
REFERENCE    2 (bases 1 to 108)
AUTHORS     Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE       Direct Submission
JOURNAL      Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
            Children's Hospital of Philadelphia, 1004F Abramson Research
            Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
            Location/Qualifiers
FEATURES     source
              1..108
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="TscAlu2"
              1..108
              /note="scAlu"
              /rpt_family="Alu"
              /rpt_type="dispersed"
              23 a 39 c 30 g 16 t
BASE COUNT   23 a 39 c 30 g 16 t
ORIGIN
Query Match      0.3%; Score 85.8; DB 11; Length 108;
Best Local Similarity 92.8%; Pred. No. 0.00045;
Matches 90; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 24035 GCCTTAATCCAGCAGCTTTGGAGCGTGAGTGGTGAATCACAGAGTCAGGAGATCAA 24094
|||||
Db 1 GCCTTAATCCAGCAGCTTTGGAGCGCGGCGGATCACAGGAGTCAGGAGATCGA 60
|||||

QY 24095 GACCATCTGGCCACATGTTCAAAACCCGCTCTAC 24131
|||||
Db 61 GACCATCTGGCTAAACAAGGTGAACCCCGTCTCTAC 97
|||||

RESULT 2
HSLDLRN2
LOCUS        Human LDL-receptor gene intron 14 fragment (normal gene).
DEFINITION
ACCESSION    X05250
VERSION      X05250.1 GI:34337
KEYWORDS     Alu repetitive sequence; low density lipoprotein receptor.
SOURCE       human
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
              Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 108)
AUTHORS     Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
TITLE       Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
JOURNAL      Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE      87161901
COMMENT      See X05252 for deletion junction
              Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES     Location/Qualifiers
              1..108
```

```
intron
BASE COUNT    28 a 23 c 39 g 18 t
ORIGIN
Query Match      0.3%; Score 85.2; DB 10; Length 108;
Best Local Similarity 87.7%; Pred. No. 0.00055;
Matches 93; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 24141 AAAAATTAGCCAGCATGGTAGCACATCCCTGTAATCCAGCTACTCAAGAGGCTGAGGC 24200
|||||
Db 3 AAAAATTAGCCAGCGCTGGTGGCAGGTGCTGTAATCCAGCTACTCGGAGGCTGAGGC 62
|||||

QY 24201 AGGGAATTGTTGAACCCGGGAGGTGGACATTCAGTGTGAGTGTGAG 24246
|||||
Db 63 AGGAGAATTGCTTGAACCCAGGAGGCAGAGGTGTCAGTGTGAGCCGAG 108
|||||

RESULT 3
HUMALCE162
LOCUS        Human carcinoma cell-derived Alu RNA transcript, clone CE162.
DEFINITION
ACCESSION    M87924
VERSION      M87924.1 GI:174871
KEYWORDS     Alu repeat.
SOURCE       Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
              Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 107)
AUTHORS     Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE       Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
JOURNAL      J. Mol. Biol. (1992) In press
FEATURES     Location/Qualifiers
              1..107
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /cell_line="NTERa2D1"
              /dev_stage="embryo"
              /sex="male"
              /tissue_type="carcinoma"
              28 a 30 c 35 g 14 t
BASE COUNT    28 a 30 c 35 g 14 t
ORIGIN
Query Match      0.3%; Score 80.6; DB 9; Length 107;
Best Local Similarity 86.4%; Pred. No. 0.0028;
Matches 89; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 729 GCGAGGAGAATCACTTGAACCTGGGAGGCAGAGGTTGCGTGGAGGAGATGGCGCCACT 788
|||||
Db 5 GCGAGAAGATGGGTGAACCCGGGAGCGGAGCTTGCAGTGAGCGAGATCGCGCCACT 64
|||||

QY 789 GCATCCAGCCTGACCAACACAGCGAGACTCTGTCTCAAAAAA 831
|||||
Db 65 GCATCCAGCCTGGCGGACAGAGCGAGACTCCGTCTCAAAAAA 107
|||||

RESULT 4
HUMALCE272
LOCUS        Human carcinoma cell-derived Alu RNA transcript, clone CE272.
DEFINITION
ACCESSION    M87899
VERSION      M87899.1 GI:174875
KEYWORDS     Alu repeat.
SOURCE       Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
              Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

REFERENCE 1 (bases 1 to 104)
AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
source
1. .104
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Ntera2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT 22 a 26 c 37 g 19 t
ORIGIN

Query Match 0.3%; Score 79.2; DB 9; Length 104;
Best Local Similarity 87.0%; Pred. No. 0.0046;
Matches 87; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 24015 GGCCTGGTGCAGTGCAGCTGCTGTAATCCAGCACTTTGGCAGCTGAGGTGGCTGAA 24074
|||||
Db 5 GGCCTGGTGCAGTGCAGCTGCTGTAATCCAGCACTTTGGCAGCTGAGGTGGCTGAA 64
|||||

QY 24075 TCACGAGGTGAGGATCAAGACCTCTGTCGCAACATGG 24114
|||||
Db 65 TTCCGAGGCCAGGATCAAGACCGCTGCTAACATGG 104
|||||

RESULT 5
LOCUS HSLDLRN2/c 108 bp DNA PRI 20-MAY-1992
DEFINITION Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION X05250
VERSION X05250.1 GI:34337
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
Williamson,R. and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901
COMMENT See X05252 for deletion junction
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.

FEATURES
source
1. .108
/organism="Homo sapiens"
/db_xref="taxon:9606"
Intron
1. .108
/note="intron XIV fragment"
BASE COUNT 28 a 23 c 39 g 18 t
ORIGIN

Query Match 0.3%; Score 77.6; DB 10; Length 108;
Best Local Similarity 82.4%; Pred. No. 0.008;
Matches 89; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 11190 CTCAACTACTGCAACCTCCCTGATTCAGGCAATTCCTGCTCAGCCCTCCGG 11249
|||||
Db 108 CTCGGCTCACTGCAACCTGCTCCTGGTTCAAGCAATTCCTGCTCAGCCCTCCGG 49
|||||

QY 11250 AGTAACTGCCACTACAGTGGCCACACACACAGTGGCTTAATTTTT 11297
|||||
Db 48 AGTAGTGGGATTACAGGCACCTGCCACACCGCTGGCTAATTTTTGT 1
|||||

RESULT 6

LOCUS HSU67804 108 bp RNA PRI 01-AUG-1997
DEFINITION Human small cytoplasmic Alu transcript.
ACCESSION U67804
VERSION U67804.1 GI:2289918
KEYWORDS Alu.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 108)
AUTHORS Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE cDNAs derived from primary and small cytoplasmic Alu (scAlu)
transcripts
JOURNAL J. Mol. Biol. 271 (2), 222-234 (1997)
MEDLINE 97415756
REFERENCE 2 (bases 1 to 108)
AUTHORS Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
Children's Hospital of Philadelphia, 1004F Abramson Research
Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

FEATURES

source
1. .108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TscAluJ"
repeat_region
1. .108
/note="scAlu"
/rpt_family="Alu"
/rpt_type="dispersed"

BASE COUNT 26 a 38 c 26 g 18 t
ORIGIN

Query Match 0.3%; Score 77.8; DB 11; Length 108;
Best Local Similarity 87.6%; Pred. No. 0.0074;
Matches 85; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 24035 GCCTGTAATCCCAGCAGCTTTGGAGGCTGAGTGGTGAATCAGAGGTGAGGATCAA 24094
|||||
Db 1 GCCTGTAATCCCAGCAGCTTTGGAGGCCAAAGCGGAGGATCACAGGTGAGGATCGA 60
|||||

QY 24095 GACCATCTTGGCCCAACATGGTGAACCCCTCTCTAC 24131
|||||
Db 61 GACCATCTTGGCTTAACATGGTGAACCCCTCTCTTC 97
|||||

RESULT 7

LOCUS HUMALCE221 103 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 103)
AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
source
1. .103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Ntera2D1"
/dev_stage="embryo"
/sex="male"

```

BASE COUNT      25 a      27 c      33 g      18 t
ORIGIN
/tissue_type="carcinoma"

Query Match      0.3%; Score 76; DB 9; Length 103;
Best Local Similarity 85.0%; Pred. No. 0.014;
Matches 85; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 24167 TCCTGTATCCAGCTACTCAAGAGGTGAGGCGAGGGAATGCTTGAACCCGGGAGGT 24226
|||||
Db 4 TCCTGTATCCAGCTACTCAAGGGAAGTAAGGCGAGGAGAAATCGTTGAACCCGGGAGGC 63
|||||

QY 24227 GGACATCTGAGTGAGCTGAGATCGCACCTACACTACCTCCAG 24266
|||||
Db 64 GGAGGTTGAGTGAGCCGAGAGATCGTGCCATTGCACTCCAG 103
|||||

RESULT 8
LOCUS HSU67808 108 bp RNA PRI 01-AUG-1997
DEFINITION Human small cytoplasmic Alu transcript.
ACCESSION U67808
VERSION U67808.1 GI:2289922
KEYWORDS Alu.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE cDNAs derived from primary and small cytoplasmic Alu (scAlu)
transcripts
JOURNAL J. Mol. Biol. 271 (2), 222-234 (1997)
MEDLINE 97415756
AUTHORS Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
Children's Hospital of Philadelphia, 1004F Abramson Research
Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES
source
Location/Qualifiers
1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TscAlu7"
/note="scAlu"
/rpt_type="dispersed"
repeat_region 1..108
BASE COUNT 22 a 28 g 21 t
ORIGIN

Query Match      0.3%; Score 76.2; DB 11; Length 108;
Best Local Similarity 86.6%; Pred. No. 0.013;
Matches 84; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 24035 GCCTGTATCCAGCACATTGGGAGCGTGAGGTGGTGAATCACAGGTCAGGAGATCAA 24094
|||||
Db 1 GCCTGTATCCAGCACATTGGGAGCGCAAGTCGGTGGATCACAAAGTCAGGAGTTTGA 60
|||||

QY 24095 GACCATCTGGCCAACTGGTGAACCCCGCTCTAC 24131
|||||
Db 61 GACCAGCTGGCCACATGGTGAAACTCGGCTTTCC 97
|||||

RESULT 9
LOCUS HS81C8R 103 bp DNA STS 05-SEP-1991
DEFINITION Human sequence tagged site 81C8R DNA from 19q13.
ACCESSION X57789
VERSION X57789.1 GI:23938

KEYWORDS STS; myotonic dystrophy.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Aldridge,P.L.
TITLE Direct Submission
JOURNAL Submitted (12-FEB-1991) F.I. Aldridge, ICI Pharmaceuticals,
Alderley Park, Macclesfield, Cheshire, SK10 4TG, UK
REFERENCE 2 (bases 1 to 103)
AUTHORS Butler,R., Riley,J.H., Ogilvie,D.J., Anand,R., Buxton,J.,
Davies,J., Johnson,K. and Markham,A.F.
TITLE Two sequence-tagged sites defining the ends of a 380 kb YAC clone
from 19q13
JOURNAL Nucleic Acids Res. 19 (17), 4787 (1991)
MEDLINE 91367697
COMMENT See also X57788 for STS 81C8L.
FEATURES
Location/Qualifiers
1..103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19q13"
/germline
/clone_lib="YAC library: ICI"
/clone="81C8"
BASE COUNT 29 a 28 c 23 g 22 t 1 others
ORIGIN

Query Match      0.3%; Score 75.4; DB 13; Length 103;
Best Local Similarity 87.2%; Pred. No. 0.017;
Matches 82; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 11305 CTGTAGATGGGTTTCGCATGTTGCCAGGTGTGCCAGGTGCTCAAACTCTCGACCTGAAGTG 11364
|||||
Db 1 CAGTAGAGATAGGTTTCACCATGTTGCCAGGTGTGCCAGGTGCTCAAGAACTCTCGACCTTAAGTG 60
|||||

QY 11365 TTCACCCACCTCGCCCTCCCAAGTCTGGATT 11398
|||||
Db 61 ATCCACCCACCTCGACCTCCCAAGATGCGNGAAT 94
|||||

RESULT 10
LOCUS HUMDLRAL/c 97 bp DNA PRI 07-JAN-1995
DEFINITION Human low density lipoprotein receptor gene, intron 4 (partial).
ACCESSION M14178
VERSION M14178.1 GI:187097
KEYWORDS low density lipoprotein receptor-1.
SEGMENT 1 of 2
SOURCE Human white blood cell DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 97)
AUTHORS Hobbs,H.H., Brown,M.S., Goldstein,J.L. and Russell,D.W.
TITLE Deletion of exon encoding cysteine-rich repeat of low density
lipoprotein receptor alters its binding specificity in a subject
with familial hypercholesterolemia
JOURNAL J. Biol. Chem. 261 (28), 13114-13120 (1986)
MEDLINE 87008518
COMMENT Analysis of the LDL-receptor gene of a patient with familial
hypercholesterolemia (FH) revealed the deletion of exon 5 resulting
from a homologous recombination between repetitive Alu sequences of
intron 4 and intron 5.
FEATURES
source
Location/Qualifiers
1..97
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19p13.3"
<1..97
/gene="LDLR"
intron
```

[illegible][illegible]

```
Query Match          0.3%; Score 74.6; DB 10; Length 108;
Best Local Similarity 81.9%; Pred. No. 0.023;
Matches 86; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 24141 AAAAATTAGCCAGGCGATGTAGCATCGCTGTATATCCAGCTACTCAAGAGGCTGAGGC 24200
|||||
Db 106 AAAAATTAGCCAGGCGTGTGGCAGGTGCCTGTATATCCAGCTACTCGGGAGGCTGAGGC 47
|||||

Qy 24201 AGGGGAATTGCTTGAACCCGGGAGGTGACATTCGAGTGAGCTGA 24245
|||||
Db 46 AGGAAATGGTTTGAACCCAGGAGGACAGGTTGTGGTGAGGCCGA 2

RESULT 14
HSLDLRD2          HSLDLRD2          108 bp          DNA          PRI          20-MAY-1992
LOCUS              Human LDL-receptor mutated gene with intron 14 deletion junction.
ACCESSION          X05251
VERSION            X05251.1 GI:34336
KEYWORDS            Alu repetitive sequence; low density lipoprotein receptor.
SOURCE              human.
ORGANISM            Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE           1 (bases 1 to 108)
AUTHORS             Horsthemke, B., Beisiegel, U., Dunning, A., Havinga, J.R.,
                    Williamson, R., and Humphries, S.
TITLE               Unequal crossing-over between two alu-repetitive DNA sequences in
                    the low-density-lipoprotein-receptor gene. A possible mechanism for
                    the defect in a patient with familial hypercholesterolaemia
JOURNAL             Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE             87161901
COMMENT              *source: hypercholesterol aemia
                    See X05250 for corresponding normal gene sequence
                    In the defective LDL-receptor gene the deletion occurred between two
                    alu-repetitive sequences, that are in the same direction, the
                    deletion eliminates exons 13 and 14 and changes the reading frame
                    of the resulting spliced mRNA.
                    Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.

FEATURES             Location/Qualifiers
     source          1..108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TscAlu6"
     intron          1..108
                     /note="intron XIV fragment"
BASE COUNT          28 a 20 c 40 g 20 t
ORIGIN

Query Match          0.3%; Score 74.6; DB 10; Length 108;
Best Local Similarity 81.9%; Pred. No. 0.023;
Matches 86; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 24141 AAAAATTAGCCAGGCGATGTAGCATCGCTGTATATCCAGCTACTCAAGAGGCTGAGGC 24200
|||||
Db 3 AAAAATTAGCCAGGCGTGTGGCAGGTGCCTGTATATCCAGCTACTCGGGAGGCTGAGGC 62
|||||

Qy 24201 AGGGGAATTGCTTGAACCCGGGAGGTGACATTCGAGTGAGCTGA 24245
|||||
Db 63 AGGAAATGGTTTGAACCCAGGAGGACAGGTTGTGGTGAGGCCGA 107
|||||

RESULT 15
HSU67807          HSU67807          110 bp          RNA          PRI          01-AUG-1997
LOCUS              Human small cytoplasmic Alu transcript.
DEFINITION          U67807
ACCESSION           U67807
VERSION             U67807.1 GI:2289921
KEYWORDS            Alu.
SOURCE              human.
```

```
ORGANISM            Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE           1 (bases 1 to 110)
AUTHORS             Shaikh, R.H., Roy, A.M., Kim, J., Batzer, M.A. and Deininger, P.L.
TITLE               cDNAs derived from primary and small cytoplasmic Alu (scAlu)
                    transcripts
JOURNAL             J. Mol. Biol. 271 (2), 222-234 (1997)
MEDLINE             97415756
REFERENCE           2 (bases 1 to 110)
AUTHORS             Shaikh, R.H., Kim, J., Batzer, M.A. and Deininger, P.L.
TITLE               Direct Submission
JOURNAL             Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
                    Children's Hospital of Philadelphia, 1004F Abramson Research
                    Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

FEATURES             Location/Qualifiers
     source          1..110
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TscAlu6"
     repeat_region   1..110
                     /note="scAlu"
                     /rpt_family="Alu"
                     /rpt_type="dispersed"
BASE COUNT          26 a 39 c 24 g 21 t
ORIGIN

Query Match          0.3%; Score 74.4; DB 11; Length 110;
Best Local Similarity 84.0%; Pred. No. 0.025;
Matches 84; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 13530 GCCTGTATATCCAGCACATTTGGGAGGCCAAGCGAGGATCCTTGAAGCCAGGAGTTC 13589
|||||
Db 1 GCCTGTATATCCAGCACATTTGAGAGGCCAAAGTGGTGTGATCCTTGAGCCAGGAGTTC 60
|||||

Qy 13590 AAGACGACGCTGCCCAACATGCGCAAAACCCCTGGCTCTACC 13629
|||||
Db 61 AAGACGACGCTGTCAAACATGCTGAAACCCCATCTTTCCC 100
|||||

Search completed: June 16, 2000, 20:09:08
Job time: 199832 sec
```



```
AC X12095;
DE 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment TIGR-A003M18a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PA (WHEED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI: 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;

Query Match 0.3%; Score 73.2; DB 1; Length 108;
Best Local Similarity 85.2%; Pred. No. 0.037;
Matches 92; Conservative 1; Mismatches 14; Indels 1; Gaps 1;

Qy 13532 CTGTATCCACGAC-TTTGGAGGCCAAGCGAGCGGATGACCTTGAAGCCAGGAGTTCA 13590
Db ||||||||||| ||||||||||| ||||||||||| ||||||||||| |||||||||||
Qy 13591 AGACCAGCCTGCCAACATGCGAAACCCCTGCTCTACCAAAATACA 13638
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 48 AGACCATCTGGCCCAACAYAGGAAACCTCTCTACAAAAAGACA 1
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
ID T24892 standard; cDNA to mRNA; 100 BP.
AC T24892;
DE 05-NOV-1996 (first entry)
DE Human gene signature HUMGS06998.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATSU) MATSUBARA K.
PA (OKUBU) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
```

```
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1720; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;

Query Match 0.2%; Score 68.8; DB 1; Length 100;
Best Local Similarity 79.8%; Pred. No. 0.14;
Matches 79; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 13704 ATCACTTCAACCGGAGCGAGGTTCCAGTCAGTCAGTCGATTCGCCACTACACTACAG 13763
Db ||| ||||||| ||||||||||| ||||||||||| ||||||||||| ||| ||||| |||
Qy 2 ATCGCTTGAACCTGGAGGCGAGATTTGCAATNAGTCAGATTCGCCCTTGCCTCCG 61
Db ||||||||||| ||||||||||| ||||||||||| ||||||||||| |||||||||||

Qy 13764 CTTGGGTGACAGAGGAGATTTCTGTCTCAAAAAAATAA 13802
Db ||||||||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||

RESULT 4
X12095
ID X12095 standard; DNA; 108 BP.
AC X12095;
DE 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment TIGR-A003M18a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PA (WHEED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI: 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;
```

```
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;

Query Match 0.2%; Score 68.4; DB 1; Length 108;
Best Local Similarity 82.4%; Pred. No. 0.16;
Matches 89; Conservative 1; Mismatches 17; Indels 1; Gaps 1;

QY 16714 TGTATTTTATAGATAGAGGTTTCAATAGTGGCCAGGCTGGTCTCAAACTCCGACC 16773
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TGTCTTTTGTAGATAGAGGTTTCTCTGTGTGGCCAGGATGGTTCGAACCTCTGAC 60

QY 16774 CTCAAGTGATCCTCCCTCGCTCGCCCTCC-ATATGCTGGGATTACAG 16820
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TTCAAGTGATCGCTCGCTGGCCCTCCCAAAAGTCTGGGATTATAG 108

RESULT 5
X12087/c
ID X12087 standard; DNA; 100 BP.
AC X12087;
DE Human biallelic polymorphic DNA fragment EST98276a.
DE Polymorphism: biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
PI WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich kidney disease, hereditary
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, Ehlers-Danlos
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match 0.2%; Score 63.4; DB 1; Length 100;
Best Local Similarity 76.8%; Pred. No. 0.76;
Matches 76; Conservative 1; Mismatches 22; Indels 0; Gaps 0;

QY 3638 GTGGCTCAAGCCTTAATCCCAACACTTTGGAGGCTTAAGTGGGAGGATTGCTTGAGCC 3697
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 99 GTGACTCACACCTATAATCTCGGCATCTTGGAGGCTTAGGAGGAGGATTGTTTGAAC 40

QY 3698 CAGTAGTTCAGACGACCGCTGGGCAACATGAGAAACCC 3736
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 39 CAGGAGCTCAAGACCATCTCTGGGAAACATAGCAAGACTC 1

RESULT 7
T24892/c
ID T24892 standard; cDNA to mRNA; 100 BP.
AC T24892;
DE Human gene signature HUMGS06998.
DE Gene signature; messenger RNA: mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
```

```
Db 39 CAGGAGCTCAAGACCAKCKCTGGGAAACATAGCAAGACTC 1

RESULT 6
X12085/c
ID X12085 standard; DNA; 100 BP.
AC X12085;
DE Human biallelic polymorphic DNA fragment EST98276c.
DE Polymorphism: biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
PI WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 218; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich kidney disease, hereditary
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, Ehlers-Danlos
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 22 A; 25 C; 22 G; 30 T;

Query Match 0.2%; Score 63.4; DB 1; Length 100;
Best Local Similarity 76.8%; Pred. No. 0.76;
Matches 76; Conservative 1; Mismatches 22; Indels 0; Gaps 0;

QY 3638 GTGGCTCAAGCCTTAATCCCAACACTTTGGAGGCTTAAGTGGGAGGATTGCTTGAGCC 3697
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 99 GTGACTCACACCTATAATCTCGGCATCTTGGAGGCTTAGGAGGAGGATTGTTTGAAC 40

QY 3698 CAGTAGTTCAGACGACCGCTGGGCAACATGAGAAACCC 3736
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 39 CAGGAGCTCAAGACCATCTCTGGGAAACATAGCAAGACTC 1

RESULT 7
T24892/c
ID T24892 standard; cDNA to mRNA; 100 BP.
AC T24892;
DE Human gene signature HUMGS06998.
DE Gene signature; messenger RNA: mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
```

PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PI (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 1720; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;

Query Match 0.2%; Score 62.4; DB 1; Length 100;
Best Local Similarity 75.8%; Pred. No. 1;
Matches 75; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 16553 TTTTTCCTGAGTGGAGTCTTACTGTGCTGCCTCAAGCTGAGTGGCGACAATC 16612
DB 100 TTTGTTGTTTCAACAGAGTGTCTACTGTGTCACCCAGGCGGAGTGAAGGTGCAATC 41

QY 16613 TCAGCTCACTGCAACCTCTGCTTCTTGGSTTCAAGCAAT 16651
DB 40 TCAGCTNATTGCAAAATTCCTGCTCCAGGTTCAAGCGAT 2

RESULT 8
ID X12086/c
AC X12086;
DE Human biallelic polymorphic DNA fragment EST98276b.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI: 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as aquagenic pruritus, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary

CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 21 A; 25 C; 22 G; 31 T;

Query Match 0.2%; Score 62.2; DB 1; Length 100;
Best Local Similarity 76.8%; Pred. No. 1.1;
Matches 76; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 3638 GTGGCTCAAGCTGTAAATCCCAACACTTTGGGAGGCTTAGGTGGGAGGATTCCTTGAGCC 3697
DB 99 GTGATCTCACACCTATAATCTTGGCACTTTAGGAGCTKAGGAAGGAGGATTTCTTGAAC 40

QY 3698 CAGTAGTTCAAGACACCGCTGGCAACATGGAGAACCC 3736
DB 39 CAGGAGCTCAAGACCATCTCTGGGAAACATAGCAAGACTC 1

RESULT 9
T25854
ID T25854 standard; cDNA to mRNA; 91 BP.
AC T25854;
DE Human gene signature HUMGS08084.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PI (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1944; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 91 BP; 18 A; 22 C; 28 G; 18 T;

Query Match 0.2%; Score 61; DB 1; Length 91;
Best Local Similarity 78.7%; Pred. No. 1.6;
Matches 70; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 13704 ATCAGTTCAACCGGGGAGGAGGAGTTCAGTGCAGTTCGATTCGCCACACACTACAG 13763
DB 2 ATCAGTTGAGCCTAGGAGGAGGAGNGGTTCAAGTGCAGTGCAGTTCGCTCCGCTCCAG 61

```
QY 13764 CCTGGGTGACAGAGAGATTCTGTCTCA 13792
      ||| ||| ||| ||| ||| ||| ||| |||
DB 62 CCTNGGTGACAGCGTGAGANNCTGTCTCA 90

RESULT 10
T20927/c
ID T20927 standard; cDNA to mRNA; 103 BP.
AC T20927.
DT 24-JUL-1996 (first entry)
DE Human gene signature HUMGS02180.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-Al.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 758-759; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 103 BP; 22 A; 27 C; 21 G; 31 T;

Query Match 0.2%; Score 61.2; DB 1; Length 103;
Best Local Similarity 75.0%; Pred. No. 1.5;
Matches 75; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 13607 CATGGCAAAACCTGCTCTACCAAAAATACACAATTAGCTGGGATTTGGGCACATGC 13666
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 100 CATGGAGAAATACTGTCCCTCTTAAACAAATACAAATACAGCTGGACATGTTGGCACAC 41

QY 13667 CTGTAATCCAGCTACTTGGGAGGCTGAAGACCAAGAATC 13706
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 40 CTGTACCACAGCTACTTGGGAGGCTGAAGTGGGAGGATC 1

RESULT 11
T26828
ID T26828 standard; cDNA to mRNA; 108 BP.
AC T26828.
DT 14-NOV-1996 (first entry)
DE Human gene signature HUMGS09078.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-Al.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
```

```
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 2182; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 18 A; 33 C; 23 G; 28 T;

Query Match 0.2%; Score 61; DB 1; Length 108;
Best Local Similarity 73.8%; Pred. No. 1.6;
Matches 76; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 28509 GATCTCTTACCTTGATCCACCCGCCCTCAGCTCCCAAGTGCAGGATTACAGGCAT 28568
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 GATCTCTTACCTTGATCCACCCGCCCTCAGCTCCCAAGTGCAGGATTACAGGCAT 60

QY 28569 GAGCCACCGTCCAGCCCTCTTTTCTTTTCTTTATTAAGACAAG 28611
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 GAGCCACCGTCCAGCCCTCTTTTCTTTTCTTTATTAAGACTGTACAGG 103

RESULT 12
T26213
ID T26213 standard; cDNA to mRNA; 103 BP.
AC T26213.
DT 13-NOV-1996 (first entry)
DE Human gene signature HUMGS08452.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-Al.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 2029; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
```


PT New nucleic acid probes - have a labelled low frequency
PT repetitive sequence for detecting overlaps among cloned DNA
PS Disclosure; Page 8; 41pp; English.
CC Genomic human placental DNA was mixed with a pair of PCR primers
CC (see Q29014 and Q29015). The amplified DNA products were separated
CC on an agarose gel and fragments in the 300-1000bp region were
CC isolated. The fragments were ligated to M13mp19 RF DNA from whic
CC the 12bp SalI-BamHI insert had been removed. Competent JM109 were
CC transformed by the ligation mixture and transformants were plated
CC on NZY plates contg. beta-galactosidase indicator dye. Duplicate
CC filter replicates were screened with two probes, one taken from the
CC internal region of an Alu repeat (see Q29016) and the other a 5kb
CC fragment contg. an L1 sequence from the region 5' to the human
CC gamma-globin gene. Phage plaques which did not hybridise to either
CC probe and did not react with the dye indicator were selected.
CC Single-stranded DNA was extracted from them to isolate low-frequency
CC repeat sequence probes LF12, LF15, LF16, LF17, LF18 and LF22.
CC See Q29013-Q29017 and Q29021-Q29038.
SQ Sequence 69 BP; 19 A; 18 C; 21 G; 11 T;

Query Match 0.2%; Score 58.4; DB 1; Length 69;
Best Local Similarity 91.2%; Pred. No. 3.5;
Matches 62; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 13544 CACTTTGGAGGCCAAGCGGAGTCACTTGAAGCCAGGAGTTCAAGACCAAGCCTGCC 13603
|||||
Db 1 CACTTTGGAGGCCAAGCGGAGTCACTTGAAGCCAGGAGTTCAAGACCAAGCCTGCC 60
|||||

Qy 13604 CACATGG 13611
|||||
Db 61 CACATGG 68

Search completed: June 16, 2000, 21:16:14
Job time: 202685 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2000, 03:46:29 ; Search time 8516.13 seconds
(without alignments)
13802.928 Million cell updates/sec

Title: US-08-852-495c-1_copy_140000_169000
Perfect score: 29001
Sequence: 1 CCTCCCAATCCCATATGCAC.....TGTATCACAGAAAGTTACC 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*

- 1: em_est1:*
- 2: em_est2:*
- 3: em_est3:*
- 4: em_est4:*
- 5: em_est5:*
- 6: em_est6:*
- 7: em_est7:*
- 8: em_est8:*
- 9: em_est9:*
- 10: em_est10:*
- 11: em_est11:*
- 12: em_est12:*
- 13: em_est13:*
- 14: em_est14:*
- 15: em_est15:*
- 16: em_est16:*
- 17: em_est17:*
- 18: em_est18:*
- 19: em_est19:*
- 20: gb_est1:*
- 21: gb_est2:*
- 22: gb_est3:*
- 23: gb_est4:*
- 24: gb_est5:*
- 25: gb_est6:*
- 26: gb_est7:*
- 27: gb_est8:*
- 28: gb_est9:*
- 29: gb_est10:*
- 30: gb_est11:*
- 31: gb_est12:*
- 32: gb_est13:*
- 33: gb_est14:*
- 34: gb_est15:*
- 35: gb_est16:*
- 36: gb_est17:*
- 37: gb_est18:*
- 38: gb_est19:*
- 39: gb_est20:*
- 40: gb_est21:*
- 41: gb_est22:*
- 42: gb_est23:*
- 43: gb_est24:*
- 44: gb_est25:*

- 45: gb_est26:*
- 46: gb_est27:*
- 47: gb_est28:*
- 48: gb_est29:*
- 49: gb_est30:*
- 50: gb_est31:*
- 51: gb_est32:*
- 52: em_est20:*
- 53: em_est21:*
- 54: em_est22:*
- 55: em_est23:*
- 56: em_est24:*
- 57: em_est25:*
- 58: em_est26:*
- 59: gb_est33:*
- 60: gb_est34:*
- 61: gb_est35:*
- 62: gb_est36:*
- 63: gb_est37:*
- 64: gb_est38:*
- 65: em_est27:*
- 66: em_est28:*
- 67: em_est29:*
- 68: em_est30:*
- 69: gb_est39:*
- 70: gb_est40:*
- 71: gb_est41:*
- 72: gb_est42:*
- 73: gb_est43:*
- 74: gb_est44:*
- 75: em_est31:*
- 76: em_est32:*
- 77: em_est33:*
- 78: em_est34:*
- 79: gb_est45:*
- 80: gb_est46:*
- 81: gb_est47:*
- 82: gb_gss1:*
- 83: gb_gss2:*
- 84: gb_gss3:*
- 85: gb_gss4:*
- 86: em_gss1:*
- 87: em_gss2:*
- 88: em_gss3:*
- 89: em_gss4:*
- 90: gb_gss5:*
- 91: gb_gss6:*
- 92: gb_gss7:*
- 93: gb_gss8:*
- 94: gb_gss9:*
- 95: em_gss5:*
- 96: em_gss6:*
- 97: em_gss7:*
- 98: em_gss8:*
- 99: em_gss9:*
- 100: em_gss10:*
- 101: em_gss11:*
- 102: gb_gss10:*
- 103: gb_gss11:*
- 104: em_gss12:*
- 105: gb_gss12:*
- 106: gb_gss13:*
- 107: gb_gss14:*
- 108: gb_gss15:*
- 109: gb_gss16:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result % Query

No.	Score	Match	Length	DB	ID	Description
C 1	93.8	0.3	109	30	AA243009	AA243009 zr25h02.s
C 2	92.2	0.3	106	37	AA703692	AA703692 ag81a10.r
C 3	91.6	0.3	106	105	AQ264176	AQ264176 CITBI-EL-
C 4	90.8	0.3	110	106	AQ386882	AQ386882 RPCI11-13
C 5	87.4	0.3	105	105	AQ282107	AQ282107 RPCI11-94
C 6	86	0.3	103	38	AA807640	AA807640 nc08b05.s
C 7	85.4	0.3	103	108	AQ535244	AQ535244 RPCI-11-3
C 8	85	0.3	109	22	H11143	H11143 ym09c06.r1
C 9	85	0.3	110	39	AA897366	AA897366 am06h02.s
C 10	85	0.3	110	94	AQ003188	AQ003188 RPCI11-ID
C 11	84.6	0.3	107	35	AA565533	AA565533 nk42b11.s
C 12	84.6	0.3	109	30	AA244173	AA244173 nc05h06.s
C 13	82.6	0.3	105	28	AA078003	AA078003 7H12D08.C
C 14	82.8	0.3	110	30	AA244245	AA244245 nc07a04.s
C 15	82.8	0.3	110	106	AQ386882	AQ386882 RPCI11-13
C 16	82.2	0.3	103	35	AA570438	AA570438 nk63g02.s
C 17	82	0.3	103	84	B48914	B48914 RPCI11-4A12
C 18	82.2	0.3	103	108	AQ534922	AQ534922 RPCI-11-3
C 19	82	0.3	106	63	AT991750	AT991750 wt48e01.x
C 20	81.8	0.3	109	94	AQ029690	AQ029690 RPCI11-41
C 21	82	0.3	110	32	AA369482	AA369482 EST80906
C 22	81.2	0.3	104	108	AQ544583	AQ544583 CITBI-EL-
C 23	81.2	0.3	105	109	AQ637292	AQ637292 RPCI-11-4
C 24	81.2	0.3	106	30	AA250812	AA250812 zs06a05.s
C 25	81.4	0.3	107	35	AA583252	AA583252 nn41e04.s
C 26	81.4	0.3	108	84	B65160	B65160 CIT-HSP-201
C 27	80.6	0.3	103	108	AQ535244	AQ535244 RPCI-11-3
C 28	80.8	0.3	106	38	AA812141	AA812141 ocy48h02.s
C 29	80.6	0.3	110	29	AA177157	AA177157 nc02g07.s
C 30	80.2	0.3	101	35	AA583697	AA583697 nn58f10.s
C 31	79.8	0.3	102	84	B48088	B48088 RPCI11-4N6
C 32	80	0.3	104	105	AQ321855	AQ321855 RPCI11-11
C 33	80.2	0.3	106	30	AA250812	AA250812 zs06a05.s
C 34	80	0.3	107	33	AA385808	AA385808 EST99495
C 35	80.2	0.3	108	32	AA370029	AA370029 EST81584
C 36	80.2	0.3	109	84	B17434	B17434 345K2.TVB.C
C 37	80.2	0.3	109	84	B17434	B17434 345K2.TVB.C
C 38	80.4	0.3	110	30	AA244245	AA244245 nc07a04.s
C 39	79.8	0.3	103	108	AQ584425	AQ584425 RPCI-11-4
C 40	79.6	0.3	106	20	T55212	T55212 yd43g11.s1
C 41	79.8	0.3	107	62	AI933497	AI933497 wm74d02.x
C 42	79.8	0.3	109	94	AQ029690	AQ029690 RPCI11-41
C 43	79.2	0.3	100	30	AA252633	AA252633 zq43g05.r
C 44	79.4	0.3	106	44	AI249096	AI249096 qh73g09.x
C 45	79.4	0.3	106	108	AQ544957	AQ544957 CITBI-EL-

ALIGNMENTS

```
RESULT 1
AA243009/c
LOCUS
DEFINITION
  zr25h02.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
  cDNA clone IMAGE:664467 3' similar to contains Alu repetitive
  element;contains element LTR1 repetitive element ;, mRNA sequence.
ACCESSION
  AA243009
VERSION
  AA243009.1 GI:1873869
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
  1 (bases 1 to 109)
  Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
  Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
  Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
  Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
  WashU-NCI human EST Project
  Unpublished (1997)
  On Dec 3, 1996 this sequence version replaced gi:1126869.
```

```
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1127 Std Error: 0.00
Seq primer: -41ml3 fwd. ST from Amersham
High quality sequence stop: 102.
Location/Qualifiers
  1. 109
  /organism="Homo sapiens"
  /db_xref="GDB:5426481"
  /db_xref="taxon:9606"
  /clone="IMAGE:664467"
  /clone_lib="Stratagene NT2 neuronal precursor 937230"
  /tissue_type="neuroepithelial cells"
  /dev_stage="Ntera-2 neuroepithelial cells"
  /lab_host="SOLR (kanamycin resistant)"
  /note="Organ: brain; Vector: pBluescript SK-; Site:1:
  EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer:
  Oligo dT. Uninduced, exponentially growing neuroepithelial
  cells (Ntera-2/cl.D1). Average insert size: 1.0 kb;
  Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG
  3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT 19 a 30 c 30 g 30 t
ORIGIN
Query Match 0.3%; Score 93.8; DB 30; Length 109;
Best Local Similarity 93.3%; Pred. No. 0.062;
Matches 98; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 24032 CACGCTTAATCCAGCACCTTTGGGAGGCTGAGGTGAATCAGGAGTACAGGAGAT 24091
|||||
Db 109 CACGCTTAATCCAGCACCTTTGGGAGGCTGAGGTGAATCAGGAGTACAGGAGAT 50
QY 24092 CAAGACCATCTCTGGCCACATGCTGAACCCCGTCTCTACTATAAA 24136
|||||
Db 49 CAAGACCATCTCTGGCCACATGCTGAACCCCGTCTCTACTATAAA 5
RESULT 2
AA703692/c
LOCUS
DEFINITION
  ag81a10.r1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone
  IMAGE:1140858 5' similar to contains Alu repetitive element; , mRNA
  sequence.
ACCESSION
  AA703692
VERSION
  AA703692.1 GI:2713610
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
  1 (bases 1 to 106)
  Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
  Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
  Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
  Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
  WashU-NCI human EST Project
  Unpublished (1997)
  On Sep 12, 1996 this sequence version replaced gi:1397630.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
```

Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 53.

FEATURES

Location/Qualifiers
1..106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1140858"
/clone_lib="Stratagene hMT neuron (#937233)"
/dev_stage="hMT neurons"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo df. Differentiated, post mitotic hMT neurons. Average insert size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3'."
19 a 29 c 29 g 29 t

BASE COUNT

Query Match 0.3%; Score 92.2; DB 37; Length 106;
Best Local Similarity 92.4%; Pred. No. 0.096; Mismatches 8; Indels 0; Gaps 0;
Matches 97; Conservative 0;

QY 24032 CACGCTGTAAATCCAGCAGCTTTGGAGGCTGAGGTGGTGAATCCAGGAGTCAGGAGAT 24091
|||||
Db 106 CAGCGTGTAAATCCAGCAGCTTTGGAGGCTGAGGCGGCAGATCAGGAGTCAGGAGAT 47
|||||

QY 24092 CAAGACCATCTCGGCCAACATGGTGAACCCGCTCTCTACTAAAA 24136
|||||
Db 46 CGAGACCATCTCGGTGAACAGCGTGAACCTCGTCTCTACTAAAA 2

RESULT 3

AQ264176 106 bp DNA GSS 27-OCT-1998
LOCUS CITBI-E1-2509A2.TF CITBI-E1 Homo sapiens genomic clone 2509A2,
DEFINITION genomic survey sequence.
ACCESSION AQ264176
VERSION AQ264176.1 GI:3792743
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 106)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.
TITLE Use of a random human BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Map Building
COMMENT Unpublished (1998)
Other_GSSs: CITBI-E1-2509A2.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
Location/Qualifiers
1..106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2509A2"
/clone_lib="CITBI-E1"
/sex="male"
/cell_type="sperm"

FEATURES

Location/Qualifiers
1..106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2509A2"
/clone_lib="CITBI-E1"
/sex="male"
/cell_type="sperm"

/note="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"
BASE COUNT 25 a 30 c 34 g 17 t
ORIGIN

Query Match 0.3%; Score 91.6; DB 105; Length 106;
Best Local Similarity 91.5%; Pred. No. 0.11; Mismatches 9; Indels 0; Gaps 0;
Matches 97; Conservative 0;

QY 24017 CCGGTGCTAGTGGCTCACCCCTGTATCCAGCAGCTTTGGAGGCTGAGGTGAATC 24076
|||||
Db 1 CCGGCGCCAGAGTCTCACCCCTGTATCCAGCAGCTTTGGAGAGCCGAGCGGGTGGATC 60
|||||

QY 24077 ACGAGTCTAGGAGATCAACACCATCTCTGGCCACATGGTGAACCC 24122
|||||
Db 61 ACGAGTCTAGGAGATCAACACCCGCTCTGGCTACATGGTGAACCC 106

RESULT 4

AQ386882/c 110 bp DNA GSS 21-MAY-1999
LOCUS RPC111-134I4.TV RPCI-11 Homo sapiens genomic clone RPCI-11-134I4,
DEFINITION genomic survey sequence.
ACCESSION AQ386882
VERSION AQ386882.1 GI:4357905
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 110)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: RPC111-134I4.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1..110
/organism="Homo sapiens"
/db_xref="GDB:7551267"
/db_xref="taxon:9606"
/clone="RPCI-11-134I4"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3 6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC Library"

FEATURES

Location/Qualifiers
1..110
/organism="Homo sapiens"
/db_xref="GDB:7551267"
/db_xref="taxon:9606"
/clone="RPCI-11-134I4"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3 6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC Library"

BASE COUNT

Query Match 0.3%; Score 90.8; DB 106; Length 110;
Best Local Similarity 89.1%; Pred. No. 0.14; Mismatches 12; Indels 0; Gaps 0;
Matches 98; Conservative 0;

QY 9524 GGTTTTCACCATGTGGCCAGGCTGGTGTGAACCTCCTGACCTCAAGTGATCCACCTGCC 9583

```
|||||
Db 110 GGGTTTACCATTGTTCCAGGCTGGTCTTGAACCTTGACCTCAAGCATCCACCTGCC 51
QY 9584 TCAGCCTCACATGTTCTGGGATTACAGCGTGAGCCACCATGCTGGCC 9633
|||||
Db 50 TCAGCCTCCCAAAGTACTTGGATTACAGCGTGAGCCACTGCTCCCGGCC 1

RESULT 5
AQ282107 105 bp DNA GSS 27-APR-1999
LOCUS RPCI11-94B21.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-94B21,
DEFINITION genomic survey sequence.
ACCESSION AQ282107
VERSION AQ282107.1 GI:3907976
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 105)
ADAMS,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
TITLE Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
JOURNAL Use of human BAC End Sequences for Sequence-Ready Map Building
COMMENT Unpublished (1998)
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1..105
/organism="Homo sapiens"
/db_xref="taxon:9606"
/db_xref="IMAGE:1255473"
/clone_lib="RPCI-11-94B21"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/notes="Vector: p77T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified p77T3
vector. Library is not normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 26 a 31 c 30 g 18 t
ORIGIN
Query Match 0.3%; Score 87.4; DB 105; Length 105;
Best Local Similarity 89.5%; Pred. No. 0.34;
Matches 94; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 24029 GCTCAGCGCTGTAATCCAGCACTTTGGAGGCTGAGGTGGTGAATCAGCAGGTCAGGA 24088
|||||
Db 1 GCTCAGCGCTGTAATCCAGCACTCTGGAGGCCAAGGTGGTGGATCAGCAGGGCATGA 60
QY 24089 GATCAGACCATCCYGGCCCAACATGGTGAACCCCGTCTCTACTA 24133
|||||
Db 61 GTACGAGACCGCTGACCAACATGGTGAACCCCGTCTCTACTA 105

RESULT 6
AA807640/c 103 bp mRNA EST 05-MAR-1998
LOCUS nx08b05.s1 NCI_CGAP_GC3 Homo sapiens cDNA clone IMAGE:1255473 3'
DEFINITION similar to contains Alu repetitive element;; mRNA sequence.
```

```
AA807640 103 bp DNA GSS 18-MAY-1999
LOCUS RPCI-11-317H22.TV RPCI-11 Homo sapiens genomic clone
DEFINITION RPCI-11-317H22, genomic survey sequence.
ACCESSION AQ535244
VERSION AQ535244.1 GI:4846934
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 103)
NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL On Jan 19, 1998 this sequence version replaced gi:2151346.
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 774 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 87.
Location/Qualifiers
1..103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI_CGAP_GC3"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/notes="Vector: p77T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified p77T3
vector. Library is not normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 19 a 27 c 30 g 27 t
ORIGIN
Query Match 0.3%; Score 86; DB 38; Length 103;
Best Local Similarity 90.2%; Pred. No. 0.5;
Matches 92; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 24031 TCACGCCCTGTAATCCAGCACTTTGGGAGGCTGAGGTGGTGAATCAGCAGGTCAGGAGA 24090
|||||
Db 103 TCACACCTGTAATCCAGCACTTTGGGAGGCGCGAGGTGAGCGGATCACAAGGTCAGGAGA 44
QY 24091 TCAAGACCATCTCGGCCAACATGGTGAACCCCGTCTCTACT 24132
|||||
Db 43 TCGAGACCATCTGCTAACACGGGTGAACCCCATCTCTACT 2

RESULT 7
AQ535244 103 bp DNA GSS 18-MAY-1999
LOCUS RPCI-11-317H22.TV RPCI-11 Homo sapiens genomic clone
DEFINITION RPCI-11-317H22, genomic survey sequence.
ACCESSION AQ535244
VERSION AQ535244.1 GI:4846934
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Hominidae; Homo.
```



```
BASE COUNT      22 a      34 c      26 g      25 t
ORIGIN

Query Match      0.3%; Score 84.6; DB 35; Length 107;
Best Local Similarity 86.9%; Pred. No. 0.72;
Matches 93; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 13657 TGGCAGCATCCCTGTATCCAGCTACTCTGGGAGGCTGAAGCACAAGATCACTTGAACCG 13716
|||||
Db 107 TGGTGTGCTCCCTGTATCCAGCTACTCTAGGAGGCTGAGGCACGAGATCACTTGAACCT 48
|||||

QY 13717 GGGAGGCAGAGGTTGCAGTGAGCTGAGATTTGCCCACTACACTACAG 13763
|||||
Db 47 GGGAGGCAGAGCTTGCAGTGAGCTGAGATTTGAGCCACTGCACCTCCAG 1

RESULT 12
AA244173/c
LOCUS      AA244173      109 bp      mRNA      EST      20-AUG-1997
DEFINITION nc05h06.s1 NCI-CGAP_Prl Homo sapiens cDNA clone IMAGE:1007291
similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION  AA244173
VERSION     AA244173.1  GI:1874876
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 109)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   On Nov 29, 1993 this sequence version replaced gi:430513.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
M.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
DNA Sequencing by: Genome Systems Inc., Greg Lennon, Ph.D.
Clone distribution: Washington University Genome Sequencing Center
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 90.
Location/Qualifiers
1..109
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1007291"
/clone_lib="NCI-CGAP_Prl"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected, histologically normal
prostate epithelial cells. Double-stranded cDNA was
ligated to EcoRI adaptors, 5 cycles of PCR applied to the
cDNA with an adaptor-specific primer, and the resulting
PCR product subcloned into pAMP10 by the
method (Life Technologies). Average insert size is 600
bp. NOTE: Not directionally cloned. This library was
constructed by David Krizman."
28 a      28 c      31 g      22 t

BASE COUNT
ORIGIN
```

```
Query Match      0.3%; Score 84.4; DB 30; Length 109;
Best Local Similarity 89.2%; Pred. No. 0.75;
Matches 91; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 9478 CCCTACACCATATGCCCGGCTAATTTTGTATTTTAGTAGACAGGGTTTCACCATGT 9537
|||||
Db 109 CACCCACACCATATGCTGGCTAATTTCTGTATTTTAGTAGAGCAGGGTTTCACCATGT 50
|||||

QY 9538 TGGCCAGGCTGGTGTGTAACCTCTGACCTCAAGTGATCCACC 9579
|||||
Db 49 TGCCAGGCTGGTGCATGAACCTCTGACCTAGGTGATCCACC 8

RESULT 13
AA078003/c
LOCUS      AA078003      105 bp      mRNA      EST      24-SEP-1999
DEFINITION 7H12D08 Chromosome 7 HeLa cDNA Library Homo sapiens cDNA clone
7H12D08, mRNA sequence.
ACCESSION  AA078003
VERSION     AA078003.1  GI:1837477
KEYWORDS   EST.
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 105)
AUTHORS   Touchman, J.W., Bouffard, G.G., Weintraub, L.A., Idol, J.R., Wang, L.,
Robbins, C.M., Nussbaum, J.C., Lovett, M. and Green, E.D.
2006 expressed-sequence tags derived from human chromosome
7-enriched cDNA libraries
Genome Res. 7 (3), 281-292 (1997)
JOURNAL   97228905
MEDLINE
COMMENT   On Apr 14, 1993 this sequence version replaced gi:693433.
Contact: Eric D. Green
Genome Technology Branch
National Human Genome Research Institute/NIH
49 Convent Dr., MSC4431, Building 49, Room 2A08, Bethesda, MD 20892
Tel: 3014020201
Fax: 3014024735
Email: egreen@nhgri.nih.gov
Plate: 12 row: D column: 08
Seq primer: -21M13 (ABI).
Location/Qualifiers
1..105
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="7H12D08"
/clone_lib="Chromosome 7 HeLa cDNA Library"
/sex="female"
/cell_line="HeLa cell line; ATCC"
/lab_host="E. coli strain DH5 alpha"
/note="Vector: pAMP10; cDNA was generated from cytoplasmic
RNA using a mixture of random DNA hexamers and oligo(dT).
From this pool of cDNA, human chromosome 7-enriched cDNA
was isolated by direct cDNA selection using chromosome 7
genomic DNA (cosmids). The resulting direct-selected cDNA
was cloned into a plasmid vector using a non-directional
uracil DNA glycosylase (UDG)-mediated cloning strategy."
20 a      33 c      23 g      29 t

BASE COUNT
ORIGIN

Query Match      0.3%; Score 82.6; DB 28; Length 105;
Best Local Similarity 86.7%; Pred. No. 1.2;
Matches 91; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 13635 TACAACAATTAGCTGGCGCATTTGGCACATGCTGTAATCCAGCTACTTTGGAGGCTGA 13694
|||||
Db 105 TACAAAATTAGCTGGCGCATTTGGCGCATGTAATCCAGCTACTTTGGAGGCTGA 46
|||||

QY 13695 AGCACAAGATATCACTTTGAACCCGGGAGGCAGAGGTTGCAGTGAGC 13739
|||||
Db 45 GACCCGAGATTTGTTGAACCCAGGAGGCAGAGGTTGCAGTGAGC 1
```

RESULT	14
AA244245	
LOCUS	110 bp mRNA EST 20-AUG-1997
DEFINITION	nc07a04.s1 NC1-CGAP_Prl Homo sapiens cDNA clone IMAGE:1007406 similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION	AA244245
VERSION	AA244245.1 GI:1875104
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 110)
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
JOURNAL	Tumor Gene Index Unpublished (1997)
COMMENT	On Jan 24, 1995 this sequence version replaced gi:634306. Contact: Robert Strausberg, Ph.D. Tel.: (301) 496-1550 Email: Robert_Strausberg@nih.gov Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuauqui, M.D., Michael Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: David B. Krizman, Ph.D. cDNA Library Arrayed by: Washington University Genome Sequencing Center DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

```

Seq primer: -41ml3 fwd. ET from Amersham
High quality sequence stop: 90.
Location/Qualifiers
1..110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1007406"
/clone_lib="NCI_CGAP_Prl"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected, histologically normal
prostate epithelial cells. Double-stranded cDNA was
ligated to EcoRI adaptors, 5 cycles of PCR applied to the
cDNA with an adaptor-specific primer, and the resulting
PCR product subcloned into pAMP10 by the UDG-cloning
method (Life Technologies). Average insert size is 600
bp. NOTE: Not directionally cloned. This library was
constructed by David Krizman."
17 a 26 c 28 g 38 t 1 others
BASE COUNT
ORIGIN

```

	Query Match	0.3%	Score 82.8;	DB 30;	Length 110;
	Best Local Similarity	87.4%;	Pred. No. 1.1;		
	Matches	90;	Conservative	0;	Mismatches 13; Indels 0; Gaps 0;
QY	16554	TTTTTTTTTTGAGATGGAGCTTACTCTGTGGCTCAAGCTGGAGTGCAGTGGCACAATCT	16613		
Db	1	TTTTTTTTTTGAGATGGAGCTTGCATCTGTGCCAGGCTGCAGTGCAGTGGCAGANTCT	60		
QY	16614	CAGCTCACTCAACCTCTGCCTTCTGGGTTCAAGCAATTCCTGC	16656		
Db	61	TGGCTCACTCAACCTCTGCCTCTCTGGGTTCAAGAGATTCCTTC	103		

RESULT 15
AQ386882

LOCUS AQ386882 110 bp DNA 21-MAY-1999
DEFINITION RPC11-134I4.tv RPC11-11 Homo sapiens genomic clone RPC11-11-134I4,
genomic survey sequence.
ACCESSION AQ386882
VERSION AQ386882
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 110)
Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
Venter,J.C.
Use of BAC End Sequences from Library RPC11-11 for Sequence-Ready
Map Building
Unpublished (1997)
Other GSSs: RPC11-134I4.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetigr.org
Clones are derived from the human BAC library RPC11-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
Research Genetics (<http://inforesgen.com>). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.

FEATURES	source	Location/Qualifiers
		1. .110
		/organism="Homo sapiens"
		/db_xref="GDB:7551267"
		/db_xref="taxon:9606"
		/clone="RPC1-11-13414"
		/clone_lib="RPC1-11"
		/sex="Male"
		/cell_type="Lymphocytes"
		/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; RPC111 Human Male BAC library"
BASE COUNT	26 a	26 c 38 g 20 t
ORIGIN		

	Query Match	0.3%	Score 82.8;	DB 106;	Length 110;
	Best Local Similarity	84.5%;	Pred. No. 1.1;		
	Matches 93;	Conservative	0;	Mismatches 17;	Indels 0;
QY 13510	GGCCAGGTGTGGCTCATGCTGTATATCCACGACACTTTGGGAGGCCAAGGCAGCGGGA	13569			
Db 1	GGCCGGGAGCAGTGGCTCACGCGCTGTATCCCAAGTACTTTGGGAGGCTGAGCGAGGTGGA	60			
QY 13570	TCACTTGAAGCCGAGGAGTTCAAGACACGACCTGCGCCACATGGCAAAACCC	13619			
Db 61	TCGCTTGAGGTCAGGAGTTCAAGACACGAGCTGGACACATGCTGAAACCC	110			

Search completed: June 16, 2000, 15:09:23
Job time: 181930 sec

